Graduate Texts in Physics

Heinz Kalt · Claus F. Klingshirn

Semiconductor Optics 1

Linear Optical Properties of Semiconductors

Fifth Edition

Springer

Graduate Texts in Physics

Series Editors

Kurt H. Becker, NYU Polytechnic School of Engineering, Brooklyn, NY, USA
Jean-Marc Di Meglio, Matière et Systèmes Complexes, Bâtiment Condorcet, Université Paris Diderot, Paris, France
Sadri Hassani, Department of Physics, Illinois State University, Normal, IL, USA
Morten Hjorth-Jensen, Department of Physics, Blindern, University of Oslo, Oslo, Norway
Bill Munro, NTT Basic Research Laboratories, Atsugi, Japan
Richard Needs, Cavendish Laboratory, University of Cambridge, Cambridge, UK
William T. Rhodes, Department of Computer and Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL, USA
Susan Scott, Australian National University, Acton, Australia
H. Eugene Stanley, Center for Polymer Studies, Physics Department, Boston University, Boston, MA, USA
Martin Stutzmann, Walter Schottky Institute, Technical University of Munich, Garching, Germany
Andreas Wipf, Institute of Theoretical Physics, Friedrich-Schiller-University Jena, Jena, Germany

Graduate Texts in Physics publishes core learning/teaching material for graduate- and advanced-level undergraduate courses on topics of current and emerging fields within physics, both pure and applied. These textbooks serve students at the MS- or PhD-level and their instructors as comprehensive sources of principles, definitions, derivations, experiments and applications (as relevant) for their mastery and teaching, respectively. International in scope and relevance, the textbooks correspond to course syllabi sufficiently to serve as required reading. Their didactic style, comprehensiveness and coverage of fundamental material also make them suitable as introductions or references for scientists entering, or requiring timely knowledge of, a research field.

More information about this series at http://www.springer.com/series/8431

Heinz Kalt · Claus F. Klingshirn

Semiconductor Optics 1

Linear Optical Properties of Semiconductors

Fifth Edition

Springer

Heinz Kalt
Institute of Applied Physics
Karlsruhe Institute of Technology
Karlsruhe, Baden-Württemberg, Germany

Claus F. Klingshirn
Institute of Applied Physics
Karlsruhe Institute of Technology
Karlsruhe, Baden-Württemberg, Germany

ISSN 1868-4513 ISSN 1868-4521 (electronic)
Graduate Texts in Physics
ISBN 978-3-030-24150-6 ISBN 978-3-030-24152-0 (eBook)
https://doi.org/10.1007/978-3-030-24152-0

Originally published as a monograph
© Springer Nature Switzerland AG 1995, 2005, 2007, 2012, 2019
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To Elisabeth, Patrick and Sebastian

Preface

Semiconductors materials are nowadays found in an uncountable number of optoelectronic devices and applications assisting our everyday life. Examples are light-emitting diodes providing lighting in houses and cars, semiconductor lasers used for the transmission of data in the Internet or of telephone calls, solar cells converting sun light into electric current, and optical sensors including the camera chip integrated in our smartphones.

Realization of such high-performance devices is only possible, if the understanding of the semiconducting properties of the employed materials and related structures is well developed. The quantum mechanical properties of electrons in semiconductors are well described by the concepts of quasi-particles, effective mass, bandstructures, etc., which we will illustrate in detail in this book. Additional quantization of quasi-particle properties has been introduced artificially by reducing the dimensionality of semiconductor structures to (quasi-)two (quantum wells), one (quantum wires), and zero dimensions (quantum dots). We shall also treat these aspects in this book in detail. Then, the quantum wells have been stacked separated by narrow barriers to form superlattices, quantum dots have been assembled to form one-, two-, or three-dimensional arrays, and photonic atoms are put together to form photonic crystals' various dimensions. All of this is nowadays used to improve or even create new optoelectronic devices.

In this sense, one has the impression that the field of semiconductor science, including optics, tends toward maturity. But there are constantly new fields of basic and applied semiconductor research coming up. For example, in the case of the Group-III Nitrides used, e.g., in Blu-ray disk players or white light-emitting diode (LEDs), material science was able to fabricate highly efficient devices, while a thorough understanding of electronic and optical properties was still missing. Optimization of the properties and in particular the development of green-emitting nitride-based laser diodes needed extensive collaborative research efforts. After the exploding interest in ZnO, one of the new stars in the semiconductor community are organic/inorganic metal-halide Perovskites. These materials show enormous potential for applications, e.g., in thin-film solar cells but still need to be developed further. Equally promising are monolayer semiconductors which open up new areas

of physics due to their outstanding properties (quasi-relativistic electrons, spin-valley coupling). Often materials are incorporated in photonic structures like optical cavities to explore new physics like polariton condensation or possible applications like single-photon emitters and quantum information technologies.

It is obvious that semiconductor optics will be an active and exciting field of research for the next decades. We are convinced that a comprehensive knowledge of semiconductor optics as presented in this book is essential for further innovations.

In this light, the *aim of this book* is to explain the optical properties of semiconductors in the visible part and the adjacent infrared and ultraviolet regions of the electromagnetic spectrum. We want to elucidate the interaction of light with various quasi-particle excitations of semiconductors like phonons, electrons, excitons, etc. This will not only allow us to understand the propagation of light in the material as well as its absorption in the bulk and its reflection and refraction at interfaces. From the knowledge of this interaction, we will learn how to reveal the intrinsic properties of semiconductors by optical spectroscopy. We will further see how emission, absorption, or modulation of light by semiconductors can be utilized for the multitude of applications mentioned above.

The idea is to evoke in the reader an *intuitive phenomenological understanding* of the physical concepts and foundations of semiconductor optics. The mathematical apparatus is thus developed as simple as possible but elaborate enough to describe the main semiconductor properties and their experimental signatures. In this spirit, this present textbook is not only suitable for graduate and postgraduate students of physics, but also for students of neighboring disciplines, such as material science and electronics. The prerequisites for the reader are an introductory or undergraduate course in general physics and some basic knowledge in atomic physics and quantum mechanics. Comprehension of solid-state physics will facilitate reading of this book, although the basic concepts will always be outlined here.

Content and structure of this textbook support the concepts of *research-oriented teaching*. The reader is provided with a solid theoretical and phenomenological background. The latter is illustrated by schematic drawings and by a multitude of experimental results mainly from optical spectroscopy. Important experimental techniques used to produce the results are described in inserts. Problems are issued to apply the teaching and learning content to realistic examples and to explore more advanced topics in the literature. We also give ample reference to original publications to motivate further reading and facilitate deeper studies of summarized content.

Finally, some comments should be made on the new appearance and structure of this book. Most important, after C. Klingshirn retired a couple of years ago it was a natural consequence that I took over the further development of this textbook. I have taught courses on “Semiconductor Optics” for many years starting in 1987 during my postdoc period at North Texas State University in Denton (Texas). These courses were based to a large extent on the tutorials I previously attended as a doctoral candidate in C. Klingshirn’s research group. After finally becoming a colleague of C. Klingshirn at Universität Karlsruhe (TH), we joined forces in our

research activities. About 30 years of close interaction in research, teaching, and academic life had a fruitful impact not only on my vocational career. Many joint research results as well as didactic concepts developed during this period of close cooperation are documented in this book.

I already mentioned the continuous evolution of semiconductor optics over the years. This has made the textbook on Semiconductor Optics to grow considerably in size up to the forth edition of 2012. A rearrangement of the content and its presentation was necessary for this new edition. The book has been split into two volumes to be able to accommodate new research topics like monolayer semiconductors, spin optoelectronics, or excitonic condensates in an appropriate way. On the other hand, consolidation of some of the previous content was necessary.

The first volume of this textbook contains the topics related to linear optical properties. We will outline the content in the introduction (Chap. 1). The second volume will cover the topics nonlinear optics, high excitation phenomena, quasi-particle interaction, coherence and relaxation dynamics of optical excitations, and new phases of such excitations like the electron–hole plasma and liquid or excitonic condensates. Phenomena like optical bistability, single-photon emission, optical gain or spin storage and manipulation will be covered as well as their applications in photonics, optoelectronics, and quantum information.

Karlsruhe, Germany Heinz Kalt
February 2019

Acknowledgements

I am most grateful to Claus Klingshirn who introduced me to the area of semiconductor optics. He taught me that a descriptive approach is essential for the successful teaching of comprehensive topics and that an intuitive understanding of physic helps progressing research.

I was lucky to have had caring mentors who promoted my career not only by providing financial support, but also by valuable advice and through teaching. Particular mentioning deserve Claus Klingshirn (formerly at J. W. Goethe Universität zu Frankfurt, Universität Kaiserslautern and Universität Karlsruhe (TH)), Art Smirl (formerly at Center for Applied Quantum Electronics, NTSU, Denton), as well as Wolfgang Rühle and Hans-Joachim Queisser (both formerly at Max Planck Institut für Festkörperforschung, Stuttgart).

I profited a lot from inspiring discussions and fruitful scientific cooperation with a multitude of colleagues. There are too many to name them all. I only want to explicitly thank Klaus Bohnert and Vadim Lyssenko who supported me with their knowledge and patience when I started into new chapters of my scientific career. I also want to thank Michael Hetterich for many years of common research and teaching in semiconductor optics.

It was fun and fulfilling to teach or supervise a host of students, doctoral researchers, and postdocs. Their brilliant ideas and sometimes demanding questions essentially formed my scientific knowledge and understanding. Many of these contributions led to well-received publications and found their mentioning in this book.

Finally, I want to acknowledge the valuable contributions of Kurt Hümmer to the appendix on group theory.

Karlsruhe, Germany — Heinz Kalt
February 2019

Contents

1 Introduction 1
1.1 Overview of Semiconductor Materials 1
1.2 Aims and Concepts 4
1.3 Outline of This Book 6
Problems 7
Reference 7

Part I Basic Solid State Optics

2 Electromagnetic Waves 11
2.1 Maxwell's Equations 11
2.2 Electromagnetic Waves in Vacuum 14
2.3 Electromagnetic Waves in Matter 17
Problems 24
References 24

3 Photons 27
3.1 Quantum Aspects of Light 27
3.2 Quantization of the Electromagnetic Field 28
3.3 Dispersion Relations 30
Problems 32
References 32

4 Density of States and Distribution Functions 33
Problems 39
References 40

5 Macroscopic Optical Properties of Solids 41
5.1 Reflection and Refraction at an Interface 41
5.1.1 Boundary Conditions 41
5.1.2 The Laws of Reflection and Refraction 44

5.1.3 Conservation Laws at an Interface . . . 47
5.1.4 Fresnel's Formulae . . . 48
5.2 Extinction, Absorption and Scattering of Light . . . 52
5.3 Transmission Through a Slab of Matter and Fabry–Perot Modes . . . 54
5.4 Optical Anisotropy . . . 58
5.4.1 Birefringence and Dichroism . . . 58
5.4.2 Optical Activity . . . 65
Problems . . . 67
References . . . 68

6 Microscopic Aspects of Weak Light-Matter Coupling . . . 69
6.1 Absorption, Stimulated and Spontaneous Emission, Virtual Excitation . . . 69
6.2 Perturbative Treatment of the Linear Interaction of Light with Matter . . . 73
Problems . . . 78
References . . . 79

7 Oscillator Model of Strong Light-Matter Coupling . . . 81
7.1 Equations of Motion and the Dielectric Function . . . 81
7.2 Corrections Due to Quantum Mechanics and Local Fields . . . 85
7.3 Spectra of the Dielectric Function and of the Complex Index of Refraction . . . 87
7.4 The Spectra of Reflection and Transmission . . . 93
7.5 Interaction of Close-Lying Resonances . . . 97
Problems . . . 99
References . . . 99

8 The Polariton Concept . . . 101
8.1 Polaritons as New Quasiparticles . . . 101
8.2 Dispersion Relation of Polaritons . . . 103
8.3 Validity Range of the Polariton Concept . . . 109
8.3.1 Common Optical Properties of Polaritons . . . 110
8.3.2 Conservation of the Photon/Polariton Momentum $\hbar k$ . . . 113
8.4 Coupled Oscillators and Polaritons with Spatial Dispersion . . . 117
8.4.1 Dielectric Function and Dispersion Relation of Polaritons with Spatial Dispersion . . . 119
8.4.2 Optical Spectra and Additional Boundary Conditions . . . 122
8.5 Real and Imaginary Parts of Wave Vector and Frequency . . . 126

8.6 Surface Polaritons . . . 127
Problems . . . 129
References . . . 131

Part II Phonons

9 The Crystal Lattice and Reciprocal Space . . . 135
9.1 Description of Periodicity in Real Space . . . 135
9.2 Reciprocal Space and Quasi-momentum . . . 138
Problems . . . 141
References . . . 142

10 Vibrations of Atoms in a Crystal Lattice . . . 143
10.1 Vibrations of a String . . . 143
10.2 Vibrations of Linear Chains . . . 145
10.3 Vibrations in Three-Dimensional Systems . . . 151
Problem . . . 152
References . . . 152

11 Quantized Lattice Vibrations . . . 153
11.1 Phonons and the Concept of Quasi-particles . . . 153
11.2 Density of States and Statistics of Phonons . . . 157
Problems . . . 160
References . . . 161

12 Phonon Polaritons . . . 163
12.1 The Phonon Stop Band . . . 163
12.2 The Dispersion of Phonon Polaritons . . . 168
12.3 Surface Phonon Polaritons . . . 174
Problems . . . 175
References . . . 176

13 Phonons in Solids of Perturbed Lattice Periodicity . . . 179
13.1 Phonons in Alloy Semiconductors . . . 179
13.2 Defects and Localized Phonon Modes . . . 181
13.3 Phonons in Superlattices and Quantum-Well Structures . . . 183
13.3.1 Back-Folded Acoustic Phonons . . . 185
13.3.2 Confined Optical Phonons . . . 186
13.3.3 Interface Phonons . . . 189
13.4 Phonons in Quantum Dots . . . 190
Problems . . . 191
References . . . 191

Part III Electron States in Semiconductors

14 Electrons in a Periodic Crystal Potential 195
14.1 Adiabatic Approximation 195
14.2 One-Electron Approximation and Bloch's Theorem 197
14.3 Metals, Semiconductors, and Insulators 201
14.4 Quasi-particle Nature of Electrons and Holes in Crystalline Solids 203
14.5 The Effective-Mass Concept 205
14.6 Electron–Phonon Interaction and Consequences of Lattice Deformation 209
Problems 213
References 214

15 Semiconductor Bandstructure 215
15.1 Basics of Bandstructure Calculations 215
15.1.1 Nearly Free Electron Model 215
15.1.2 Tight-Binding Models 219
15.1.3 $k \cdot p$ Perturbation Theory 223
15.2 Bandstructure of Real Semiconductors 225
15.2.1 Semiconductors with Diamond or Zincblende Structure 225
15.2.2 Semiconductors with Wurtzite Structure 230
15.2.3 Further Semiconductors Important for Applications . . . 231
Problems 233
References 234

16 Optical Band-to-Band Transitions 235
16.1 Perturbation Theory of Direct Band-to-Band Transitions 235
16.2 Optical Functions of Semiconductors and Their Critical Points 239
16.3 Occupation of Electronic Band States 245
Problems 248
References 249

17 Low-Dimensional Semiconductor Structures 251
17.1 Density of States and Critical Points in Various Dimensions 251
17.2 Quantum Wells and Superlattices 253
17.3 Monolayer Semiconductors 262
17.4 Quantum Wires and Nanorods 266
17.5 Quantum Dots 268
Problems 269
References 270

18 Electronic Defects and Disorder 273
18.1 Defects, Defect States and Doping 273
18.1.1 Donors and Acceptors in Bulk Semiconductors 274
18.1.2 Shallow Defect Related Radiative Transitions 277
18.1.3 Deep Defects and Related Radiative Transitions 279
18.1.4 Defects and Doping in Quantum Wells 280
18.2 Disordered Systems and Localization 281
18.2.1 Anderson Model of Localization 281
18.2.2 Realizations of Disorder in Semiconductors 284
18.2.3 Weak Localization and Percolation 290
Problems 293
References 293

19 Excitations in One-Component Carrier Gases 297
19.1 Plasmons and Pair Excitations 297
19.2 Surface Plasmons 302
19.3 Plasmon-Phonon Mixed States 303
19.4 Plasmons in Low-Dimensional Semiconductors 307
19.5 Correlation Effects in Degenerately Doped Semiconductors 308
Problems 310
References 310

Part IV Excitons

20 Excitonic Quasi-particles 315
20.1 Electron–Hole Pair Excitations 315
20.2 Wannier and Frenkel Excitons 317
20.3 Corrections to the Simple Exciton Model 321
20.4 Biexcitons and Trions 325
Problems 325
References 326

21 Optical Properties of Intrinsic Excitons in Bulk Semiconductors 329
21.1 Excitons with Strong Oscillator Strength 329
21.1.1 Exciton–Photon Coupling 330
21.1.2 Consequences of Spatial Dispersion 332
21.1.3 Signature of Excitons in Optical Spectra 333
21.1.4 Spectroscopy in Momentum Space 354
21.1.5 Surface-Exciton Polaritons 361
21.1.6 Excitons in Organic Semiconductors, in Organic/Inorganic Hybrids and in Insulators 363
21.1.7 Saddle-Point and Core Excitons 366
21.2 Forbidden Exciton Transitions 368

21.2.1 Direct-Gap Semiconductors . . . 369
21.2.2 Indirect-Gap Semiconductors . . . 374
21.3 Intraexcitonic Transitions . . . 377
Problems . . . 380
References . . . 380

22 Optical Properties of Bound and Localized Excitons . . . 387
22.1 Bound-Exciton and Multi-exciton Complexes . . . 387
22.1.1 Binding of Excitons to Defects . . . 387
22.1.2 Spectroscopy of Bound Exciton Complexes . . . 389
22.2 Excitons in Disordered Systems . . . 397
Problems . . . 402
References . . . 402

23 Excitons in Low-Dimensional Semiconductor Structures . . . 405
23.1 Excitons in (Multiple) Quantum Wells . . . 405
23.1.1 Excitonic States in (Quasi-)Two-Dimensional Structures . . . 405
23.1.2 Optical Spectroscopy of Excitons in (Multiple) Quantum Wells . . . 408
23.2 Excitons in Coupled Quantum Wells and Superlattices . . . 418
23.3 Excitons in Monolayer Semiconductors . . . 427
23.4 Excitons in Quantum Wires and Nanorods . . . 432
23.5 Excitons in Quantum Dots . . . 436
23.5.1 Excitonic States in Quasi-0D Structures . . . 436
23.5.2 Spectroscopy of Optical Excitations in Quasi-0D Structures . . . 438
Problems . . . 447
References . . . 448

24 Excitons Under the Influence of External or Internal Fields . . . 453
24.1 Excitons and Carriers in Magnetic Fields . . . 454
24.1.1 Influence of Magnetic Fields on Excitonic Resonances . . . 454
24.1.2 Magnetic Fields Applied to Non-magnetic Bulk Semiconductors . . . 456
24.1.3 Magnetic Fields Applied to Bulk Diluted Magnetic Semiconductors . . . 461
24.1.4 Magnetic Fields Applied to Low-Dimensional Structures . . . 464
24.2 Excitons and Carriers in Electric Fields . . . 468
24.2.1 Electric Fields Applied to Bulk Semiconductors . . . 469
24.2.2 Electric Fields Applied to Low-Dimensional Structures . . . 472

24.3 Influence of Strain Fields on Electronic and Excitonic States 474
24.3.1 Strain Fields in Bulk Semiconductors 474
24.3.2 Strain in Structures of Reduced Dimensionality 477
Problems 479
References 479

25 Review of the Linear Optical Properties 485
Problem 487
References 488

Appendix A: Kramers–Kronig Relations 491

Appendix B: Group Theory in Semiconductor Optics 497

Appendix C: Important Data of Semiconductors 531

Appendix D: Growth and Structural Properties of Low-Dimensional Semiconductor Structures 535

Index 549

Chapter 1
Introduction

Abstract In this introductory chapter we first give a compact overview of semiconducting materials and some of their basic properties. We will then define the fundamental aims and concepts of this book and outline its contents.

1.1 Overview of Semiconductor Materials

Semiconductors are solids that have at a temperature $T = 0\,\mathrm{K}$ a series of completely filled valence bands. The uppermost band is separated by an energy gap $E_\mathrm{g} \lesssim 4\,\mathrm{eV}$ from the empty conduction bands. Semiconductors are thus insulators at this temperature. At elevated temperatures and in particular when doped these solids turn into reasonably good conductors.

Most important is, that this *conductivity* can be strongly modified, often over many orders of magnitudes. This can be achieved not only by adjusting material properties (e.g., by doping) but also by external influence like applied fields, pressure, temperature etc.. Semiconductors thus find extensive applications as current switches and amplifiers (diodes, transistors) in modern electronics but also as sensors. Finally the presence of a *bandgap* allows in most of the common semiconductors for optical generation and recombination of *electron–hole pairs*. Typical applications relying on such optical transitions are solar cells, light sensors, light-emitting diodes and laser diodes.

Presently more than 600 *element* and *compound semiconductors* are known. They are listed with their properties in the volumes of Landolt–Börnstein [82L1].

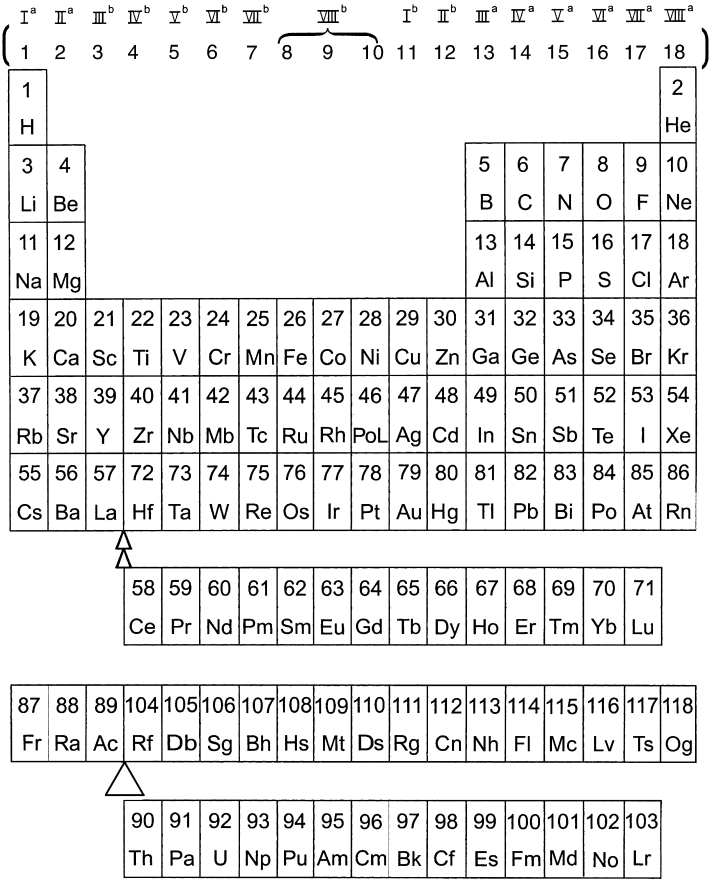
With the help of the *periodic table of the elements* (Table 1.1) we aim at identifying the most important semiconductors. In the first two lines we give the current internationally recommended numbering system for the columns used by chemists, which runs from 1 to 18, and the old one, which is frequently used by semiconductor physicists. It runs from I to VIII and has sub-classifications such as II^a and II^b.

The technically most important semiconductor is Silicon (Si). It is found in column IV^a. Silicon crystallizes in the diamond structure whose symmetry is described by

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_1

Table 1.1 Periodic table of the elements (as of 2016)



the point group O_h. The binding is completely covalent and E_g is around 1.1 eV at room temperature.

Also situated in column IV^a is carbon whose modification diamond is a semiconductor as well as Germanium (Ge). Tin (Sn) is usually a metal, but there exists a modification, which is a semi-metal namely grey tin, which also crystallizes in the diamond structure. Lead (Pb), finally, is a metal. Here we already notice a general trend, namely that the width of the bandgap decreases for elements further down the columns.

SiC also belongs to the group-IV semiconductors. SiC is notorious for forming many different crystalline polytypes. The indirect band gaps of all of them are around 3 eV, i.e., they lie between the gaps of diamond and Si.

Carbon (C) has further modifications namely graphite, graphene, carbon nanotubes, and fullerene (C_{60}). Graphite is a semi-metal with rather strong covalent sp^2 hybrid binding within the hexagonal layers and weak van der Waals binding between the layers. C_{60} is a semiconductor with $E_g \approx 2.2$ eV. Graphene and carbon nanotubes can be semiconductors, semimetals or metals.

In the diamond structure every atom is tetrahedrally surrounded by its nearest neighbors (see Fig. 9.2). We can now replace, e.g., the Ge atoms on one sub-lattice by Ga, which has one outer electron less, and the others by As, which has one electron more than Ge. This leaves the total number of electrons per unit unchanged, but adds a little bit of ionic binding to the still dominant covalent one. This procedure leads to the so-called zincblende type crystal structure with point group T_d and to the group of III–V semiconductors. To this group of materials belong the compounds of B, Al, Ga or In with N, P, As or Sb. There is again the general trend of the bandgaps within the columns of the periodic table with BN being an insulator, AlN, GaN being wide-gap semiconductors while InSb is a narrow-gap semiconductor.

The group-III nitrides actually crystallize preferentially in the hexagonal wurtzite-type structure (point group C_{6v}). In this case every atom of one type is still tetrahedrally surrounded by the others, but the arrangement of the next-nearest neighbours is such, that a hexagonal structure evolves (see Fig. 9.2).

If we repeat the step, which has led from the group IV semiconductors to the III–V compounds, once more or even twice more, we come to the II–VI (more precisely II^b–VI^a) and I–VII (more precisely I^b–VII^a) semiconductors, with increasing and finally dominating ionic binding, but still generally tetrahedral coordination. The II^a–VI^a and I^a–VII^a components are usually insulators and crystallize frequently in the rock-salt and CsCl structures.

The II–VI semiconductors comprise the compounds of Zn, Cd or Hg with O, S, Se or Te. The gap generally decreases again when going down the columns: ZnS, ZnO and CdS are wide-gap semiconductors, while the mercury (Hg) compounds are usually semimetals. The II–VI semiconductors typically crystallize in zincblende or wurtzite type structures with a few exceptions. Both structures are partly possible with only minor energetic differences, as is the case for ZnS.

Some compounds have other structures like CdO (rocksalt structure) or HgO (rhombohedral). HgS is a semi-metal in the zincblende structure but a semiconductor with a gap around 2.2 eV in the trigonal modification (red cinnabar).

The most investigated I–VII compounds are the Cu-halides and the Ag-halides, with the exception of the fluorides. Not much is known about the fluorides or the Au^+ halides concerning their properties as semiconductors.

Until now the list included only elements or binary compounds. In a similar way as above one can come to *ternary semiconductors* like $CuGaSe_2$ or even to *quaternary* ones like Ag_2CdGeS_4. Even organic molecules can be incorporated like in the case of the organic/inorganic metal halides, e.g., $CH_3NH_3PbX_3$ (with X = Cl,I,Br), and related hybrid perovskites.

Furthermore, some of the elements and many of the binary compounds form alloys partly even without a miscibility gap like $Si_{1-x}Ge_x$, $Ga_{1-y}Al_yAs$, $CdS_{1-x}Se_x$, $ZnSe_{1-x}Te_x$, $Cd_{1-y}Hg_yTe$ or $CuIn_{1-x}Ga_xSe_2$. These alloys are still, in principle, nice

crystals with a periodic lattice structure, but the lattice sites of one sub-lattice are randomly occupied by the two different atoms ($Si_{1-x}Ge_x$), anions (e.g., $CdS_{1-x}Se_x$) or cations (e.g., $Ga_{1-y}Al_yAs$). However, on a microscopic scale, the concentration fluctuations of the composition x introduce some disorder which affects the electronic and optical properties. Some alloys tend to form ordered structures for compositions close to 0.5 like $Ga_{0.5}In_{0.5}P$ adopting the so-called CuPt structure. Alloying is also possible on both sub-lattices like in $Ga_{1-y}In_yN_xAs_{1-x}$.

Most of the discussions in this book will focus on these more common semiconductors. But, there are many more, some of which we want to mention here.

The IV–VI compounds (also known as lead salts) include the compounds of Pb or Sn with S, Se and Te. They serve partly as IR laser diodes. PbSe has e.g. a narrow direct gap of 0.3 eV.

There are further elemental semiconductors like S, Se, Te and some modifications of P or I (As and Sb are considered semi-metals). There exist various oxides as semiconductors apart from the group II^B oxides like GeO_2, SnO_2 (SiO_2 called quartz is an insulator); Cu_2O; TiO_2 in its various modifications (anatase, rutil, brookit) or the highly poisonous Tl–halides.

To conclude this overview we mention *organic semiconductors* like crystals of anthracene ($C_{14}H_{10}$), pentacene ($C_{22}H_{14}$), dibenzothiophene $C_{12}H_8S$ or hexathiophene. Organic semiconductors do not fall within the focus of this book, but we will occasionally give examples of their optical properties.

1.2 Aims and Concepts

The central aim of this book now is to develop common rules and concepts which describe the electronic and in particular the optical properties of the above reviewed semiconductor materials in the linear-optics regime.

We want to describe first the macroscopic optical properties of semiconductors like transmission and reflection which are directly accessible in experiments (which in its simplest form only need a light source and our eyes to observe them). A thorough understanding of these properties require microscopic models how light interacts with matter. A simple classical model is the Lorentz-oscillator which can provide an intuitive approach to comprehend semiconductor optics with only small corrections to account for the quantum-mechanical nature of matter. The dielectric function derived in this approach nicely describes the mentioned experimental observations.

The microscopic, quantum-mechanical description of the solid uses the concept of quasi-particles and their dispersion $\omega(\boldsymbol{k})$. Examples of such quasi-particles are phonons, electrons or holes in solids, excitons, or plasmons. This concept treats the thermally or optically excited states of the solid as an ensembles of single particles moving in effective potentials which include the (mostly periodic) arrangement of the atoms in the solid and interactions with other particles or with fields. A plasmon for example is a collective excitation of the sea of free electrons in a metal (or in a highly-doped semiconductor) which moves like a particle and obeys conservation laws. The

motion of quasi-particles is characterized by its dispersion. Here the bandstructure, which is the dispersion of a single electron in the effective potential of the solid, is a prominent example.

Interaction of light with semiconductor samples can then be described as a coupling of the electromagnetic wave to polarization waves associated to the quasi-particles. Photon and quasi-particle form a new eigenstate of the solid - the polariton. The polariton model is able to describe the propagation of light in matter. The disruption of the coherent (strong) coupling of photon and quasi-particle, e.g., by scattering events leads to optical transitions between the quasi-particle states. In the limit of weak coupling (relevant for the optical transitions between electronic band states) the light is treated as a perturbation inducing transitions like absorption and stimulated emission.

An important quasi-particle necessary to describe the linear optical properties of many semiconductors is the exciton. Absorption or emission of photons close to the bandgap energy of a semiconductor involves pairs of electrons and holes. The bound states of electron–hole pairs resulting from the Coulomb interaction are called excitons. They modify in a significant way the optical spectra of semiconductors close to the absorption edge and are thus essential to understand the properties of many opto-electronic devices. The latter often utilize semiconductor nano-structures to exploit quantization effects. Excitonic effects are even more pronounced in such low-dimensional semiconductor structures. For these reasons will treat the optical properties of low-dimensional systems in detail.

We already mentioned the Lorentz-oscillator model as an intuitive approach to the macroscopic optical properties. The more elaborate theoretical treatment in terms of the semiconductor Bloch equations will be introduced in the second volume of this book. But also in this volume we will of course outline the main features of theoretical models when it is required. Examples are $\boldsymbol{k} \cdot \boldsymbol{p}$-theory for the description of the bandstructure close to the bandgap, time-dependent perturbation theory to calculate optical band-to-band transitons or group theory to describe band symmetries and optical selection rules.

An indispensable tool to illustrate optical properties is the presentation of optical spectra or of experimentally determined dispersion curves. We will give here a manifold of experimental data to support the theoretical or phenomenological concepts. But we will also demonstrate how the macroscopic and microscopic properties of semiconductors can be characterized by optical methods. To this end we will introduce the required experimental techniques like various methods of optical spectroscopy.

As already stated in the preface, this book is split into two volumes. The second volume will be devoted to semiconductor optics in the nonlinear optical/high-excitation regime, to the coherent and incoherent dynamics of optical excitations, as well as to the basics of opto-electronic, photonic and quantum-optic applications.

1.3 Outline of This Book

We summarize in *Part I* of this volume the basic concepts of solid-state optics. We will start in Chap. 2 with the description of light propagation in vacuum and matter based on Maxwell's equations. Chapter 3 introduces the quantization of the light field into photons and their dispersion relation. The general concepts of the density of states and distribution functions describing its occupation are topic of Chap. 4.

General macroscopic and microscopic approaches to optical properties of solids are presented in the next chapters. The reflection and refraction of light at the air-solid interface including the case of a slab of matter, light extinction during propagation in matter and the implications of optical anisotropy are summarized in Chap. 5. The perturbative treatment of light-matter interaction in the weak-coupling regime is outlined in Chap. 6.

Chapter 7 introduces the Lorentz oscillator model of strong light-matter coupling including the formulation of the dielectric function and the derived optical spectra within this approach. This is followed by the presentation of the basic concepts of the polariton model in Chap. 8. Its quasi-particle properties, dispersion relation and the resulting optical properties of a model solid are explained as well as the implication of coupling of the Lorentz oscillators (spatial dispersion) or propagation along an interface (surface polaritons) are described.

The second part (*Part II*) of this book called "Phonons" starts with the description and consequences of periodic crystal potentials and the concepts of reciprocal space and quasi-momentum (Chap. 9). We illustrate the formation of lattice vibrations (Chap. 10) and explain the basic properties of quasi-particles for quantized lattice vibrations, i.e., phonons in Chap. 11.

Optical phonons can couple to light and are then an ideal system to illustrate how optical properties of solids can be described in the polariton picture (Chap. 12). In the same chapter the dispersion of phonon and surface-phonon polaritons as well as experimental methods to determine them will be discussed. Chapter 13 describes phonons in semiconductors which are disordered or of reduced dimension.

In *Part III* we focus on electron states in semiconductors. We illustrate the one-electron approximation, the quasi-particle nature of electrons and holes and introduce the effective-mass concept in Chap. 14. In the next chapter (Chap. 15) we present the electronic band structure, the $\boldsymbol{k} \cdot \boldsymbol{p}$-method to deduce its basic properties at critical points like the bandgap and present the detailed band structures of the most important semiconductors. In Chap. 16 we will explain how optical transitions between bands evolve as result of the perturbation by a light field, how they determine the optical properties of the solid and how occupation of band states is described.

The quantization of electronic states in low-dimensional semiconductor structures like quantum wells, wires or dots is the topic of Chap. 17. Defect states and electrons in disordered semiconductors are treated in Chap. 18. The part on electrons closes in Chap. 19 with plasmonic excitations and correlation effects in highly-doped semiconductors.

Exciton are treated in detail in *Part IV*. We introduce the quasi-particle properties of excitons in Chap. 20. This is followed in Chap. 21 by a review of optical properties of various types of semiconductors evoked by exciton including experimental techniques for exciton spectroscopy. Excitons bound to defects or spatially localized in disordered semiconductors are the focus of Chap. 22.

Excitonic effects are especially important in low-dimensional semiconductors. The consequences of exciton confinement in quantum wells, superlatices, quantum wires and dots as well as monolayer semiconductors are elaborated in Chap. 23. The modifications of electronic states and exciton-related optical properties under the influence of magnetic, electric and strain fields are discussed in Chap. 24.

The part on excitons and also this first volume closes with a schematic summary of linear optical properties of a model semiconductor over the spectral range from IR to near-UV and in particular close to its bandgap (Chap. 25).

We add some appendices in *Part V*. Various versions of Kramers–Kronig relations are discussed in Appendix A. In Appendix B we give a compact introduction into group theory and, in particular, how it is applied to the properties of semiconductor bandstructures and to semiconductor optics. Important data of widely used semiconductors are listed in Appendix C. A short introduction into growth methods and structural properties of semiconductors an their nano-structures is given in Appendix D.

Problems

1.1 What are the basic conservation laws in nature?

1.2 Recapitulate some of the basic concepts of quantum mechanics:

- What is the Hamiltonian in classical and in quantum mechanics?
- Write down the time-independent and the time-dependent Schrödinger equation for a single particle.
- What are the eigenenergies and eigenfunctions of a one-dimensional harmonic oscillator, of the hydrogen atom, and of a particle in a square potential?

Reference

[82L1] O. Madelung, U. Rössler (ed.), *Landolt–Börnstein*. New Series, Group III, vol. 17a to i, 22a and b, 41A to D (Springer, Berlin, 1982–2001)

Part I
Basic Solid State Optics

Part I of this book on semiconductor optics is intended to summarize some basic macroscopic optical properties of solids and to give a simple microscopic model for a phenomenological understanding of these properties. We will start with a derivation of the wave equation for electromagnetic waves based on Maxwell's equations. We will discuss the propagation of light in vacuum as well as in matter and will outline the necessary physical parameters and optical functions for its description. This is followed by a short introduction into quantum aspects of light and the dispersion of photons in vacuum. The general concepts of a density of states and of distribution functions describing the occupation of theses states are introduced next. In the central chapter on macroscopic optical properties we will consider the interface between two media and—based on boundary conditions and conservation laws—the resulting consequences for light propagation. We deduce the laws of reflection and refraction and define the quantities absorption and extinction. Since semiconductors samples often have the form of plane-parallel slabs, we consider the formation of Fabry-Perot modes in such entities. Typical for solid-state materials is also optical anisotropy leading us to consider birefringence, dichroism and optical activity. We will then first introduce weak light-matter coupling and its treatment in perturbation theory. This is followed by the concept of strong coupling which we illustrate for a phenomenological model substance composed of Lorentz oscillators. Starting with the dielectric function we will derive the optical functions of this model semiconductor. Finally we will introduce the quasi-particle polariton, which is the quantized state of the mixed mode between electromagnetic and accompanying polarization wave in matter. The polariton dispersion is a useful concept to characterize light propagation in media.

Chapter 2
Electromagnetic Waves

Abstract In this chapter we derive the properties of electromagnetic waves. We start from the differential form of the macroscopic Maxwell's equations and deduce the wave equation for the electric field in vacuum. For its most frequently used solution to describe light - the plane harmonic wave - we formulate the essential properties of light waves, the relations between the oscillating electromagnetic fields, and the propagation parameters. Then we include the linear polarizability of non-magnetic, dielectric media to arrive at the wave equation in matter. By use of the complex quantities dielectric function, refractive index, and wave vector we elaborate the propagation of light in matter including its damping by extinction. We will also shortly review the properties of meta-materials, in particular the consequences of negative or zero refractive indizes.

2.1 Maxwell's Equations

Maxwell's equations (see [65M1] and [07B1, 10C1] for a modern discussion) together with the equation for the Lorentz force yield the theoretical basis of classical electro-dynamics. These equations can be written in different ways. We use here the macroscopic Maxwell's equations in their differential form.[1] These equations are given in their general form in (2.1a–f), where bold characters symbolize vectors and normal characters scalar quantities.

$$\nabla \cdot \boldsymbol{D} = \rho\,, \qquad \nabla \cdot \boldsymbol{B} = 0\,, \tag{2.1a,b}$$

$$\nabla \times \boldsymbol{E} = -\dot{\boldsymbol{B}}\,, \qquad \nabla \times \boldsymbol{H} = \boldsymbol{j} + \dot{\boldsymbol{D}}\,, \tag{2.1c,d}$$

$$\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P}\,, \qquad \boldsymbol{B} = \mu_0 \boldsymbol{H} + \boldsymbol{M}\,. \tag{2.1e,f}$$

The various symbols have the following meanings and units:

$\boldsymbol{E}$ = *electric field strength*; $1\,\mathrm{V/m} = 1\,\mathrm{m\,kg\,s^{-3}\,A^{-1}} = 1\,\mathrm{N/A\,s}$
$\boldsymbol{D}$ = *electric displacement*; $1\,\mathrm{A\,s/m^2} = 1\,\mathrm{C/m^2}$

[1] Throughout this book the internationally recommended system of units known as SI (systéme international) is used.

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_2

$\boldsymbol{H}$ = *magnetic field strength*; 1 A/m
$\boldsymbol{B}$ = *magnetic induction* or *magnetic flux density*; $1\,\mathrm{V\,s/m^2} = 1\,\mathrm{T} = 1\,\mathrm{Wb/m^2}$
ρ = *charge density*; $1\,\mathrm{A\,s/m^3} = 1\,\mathrm{C/m^3}$
$\boldsymbol{j}$ = *electrical current density*; $1\,\mathrm{A/m^2}$
$\boldsymbol{P}$ = *polarization density* of a medium, i.e., electric dipole moment per unit volume; $1\,\mathrm{A\,s/m^2}$
$\boldsymbol{M}$ = *magnetization density* of the medium, i.e., magnetic dipole moment per unit volume[2]; $1\,\mathrm{Vs/m^2}$
$\varepsilon_0 \simeq 8.859 \times 10^{-12}\,\mathrm{A\,s/V\,m}$ is the *permittivity of vacuum*
$\mu_0 = 4\pi \times 10^{-7}\,\mathrm{V\,s/A\,m}$ is the *permeability of vacuum*
$\boldsymbol{\nabla}$ = *Nabla-operator*, in Cartesian coordinates $\boldsymbol{\nabla} = (\partial/\partial x, \partial/\partial y, \partial/\partial z)$
$\dot{} = \partial/\partial t$; i.e., a dot means differentiation with respect to time.

The applications of ∇ to scalar or vector fields are usually denoted by

$$\boldsymbol{\nabla} f(\boldsymbol{r}) = \mathrm{grad\,f}(\boldsymbol{r})\ , \tag{2.2a}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{A}(\boldsymbol{r}) = \mathrm{div}\ \boldsymbol{A}(\boldsymbol{r})\ , \tag{2.2b}$$

$$\boldsymbol{\nabla} \times \boldsymbol{A}(\boldsymbol{r}) = \mathrm{curl}\ \boldsymbol{A}(\boldsymbol{r})\ , \tag{2.2c}$$

where $\boldsymbol{r}$ is the position vector.

Later we will also need the *Laplace operator* Δ, which is defined as

$$\Delta \equiv \nabla^2\ . \tag{2.3}$$

If Δ is applied to a scalar field ρ we obtain in Cartesian coordinates

$$\Delta\,\rho = \frac{\partial^2 \rho}{\partial x^2} + \frac{\partial^2 \rho}{\partial y^2} + \frac{\partial^2 \rho}{\partial z^2} \tag{2.4}$$

Application to a vector field $\boldsymbol{E}$ results in

$$\Delta \boldsymbol{E} = \begin{pmatrix} \dfrac{\partial^2 E_x}{\partial x^2} + \dfrac{\partial^2 E_x}{\partial y^2} + \dfrac{\partial^2 E_x}{\partial z^2} \\ \dfrac{\partial^2 E_y}{\partial x^2} + \dfrac{\partial^2 E_y}{\partial y^2} + \dfrac{\partial^2 E_y}{\partial z^2} \\ \dfrac{\partial^2 E_z}{\partial x^2} + \dfrac{\partial^2 E_z}{\partial y^2} + \dfrac{\partial^2 E_z}{\partial z^2} \end{pmatrix}\ . \tag{2.5}$$

Further rules for the use of ∇ and of Δ and their representations in other than Cartesian coordinates (polar or cylindrical coordinates) are found in compilations of mathematical formulae [91B1, 92S1] or [07B1].

[2] Some authors prefer to use $\boldsymbol{M}' = \boldsymbol{M}\mu_0^{-1}$ and thus $\boldsymbol{B} = \mu_0(\boldsymbol{H} + \boldsymbol{M}')$. We prefer (2.1f) for symmetry reasons.

To complete the comprehensive formulation of classical electromagnetism, we have to add the equation for the *Lorentz force* $\boldsymbol{F}$ exerted by the electric and magnetic fields on a charge q moving with a velocity $\boldsymbol{v}$

$$\boldsymbol{F} = q\boldsymbol{E} + q\boldsymbol{v} \times \boldsymbol{B} \tag{2.1g}$$

Equations 2.1a,b show that external (also called "free") electric charges comprising the charge density ρ are the source of the electric displacement and that the magnetic induction is source-free. Equations 2.1c,d demonstrate how temporally varying magnetic and electric fields generate each other. In addition, the $\boldsymbol{H}$ field can be created by a macroscopic current density $\boldsymbol{j}$. Equations 2.1e,f are the *material equations* in their general form. From them we learn that the electric displacement is given by the sum of electric field and polarization, while the magnetic flux density is given by the sum of magnetic field and magnetization. Some authors prefer not to differentiate between $\boldsymbol{H}$ and $\boldsymbol{B}$. This leads to difficulties, as can be easily seen from the fact that $\boldsymbol{B}$ is source-free (2.1b) but $\boldsymbol{H}$ is not, as follows from the inspection of the fields of every simple permanent magnet.

By applying $\boldsymbol{\nabla}\cdot$ to (2.1d) we obtain the *continuity equation* for the electric charges

$$\boldsymbol{\nabla} \cdot \boldsymbol{j} = -\frac{\partial}{\partial t}\rho\,, \tag{2.6}$$

which corresponds to the conservation law of the electric charge in a closed system.

The integral forms of (2.1) can be obtained from the differential forms by integration and the use of the theorems of Gauss or Stokes resulting in

$$\int \rho(\boldsymbol{r})\mathrm{d}V = \oint \boldsymbol{D} \cdot \mathrm{d}\boldsymbol{f} \tag{2.7a}$$

$$-\frac{\partial}{\partial t}\int \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{f} = \oint \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{s} \tag{2.7b}$$

where $\mathrm{d}V$, $\mathrm{d}\boldsymbol{f}$ and $\mathrm{d}\boldsymbol{s}$ give infinitesimal elements of volume, surface or area, and line, respectively.

In their microscopic form, Maxwell's equations contain all charges as sources of the electric field $\boldsymbol{E}_{\mathrm{micro}}$ including all electrons, protons bound in atoms as ρ_{bound} and not only the external space charges ρ. By analogy, not only the microscopic current density $\boldsymbol{j}$ has to be used as a source of $\boldsymbol{H}_{\mathrm{micro}}$ but all spins with their magnetic moments and all $l \neq 0$ orbits of charged particles have to be included as "bound" current density $\boldsymbol{j}_{\mathrm{bound}}$. The transition to macroscopic quantities can then be performed by averaging over small volumes (larger than an atom but smaller than the wavelength of light) and replacing ρ_{bound} by $-\boldsymbol{\nabla}\cdot\boldsymbol{P}$ and $\boldsymbol{j}_{\mathrm{bound}}$ by $\dot{\boldsymbol{P}} + \mathrm{curl}\,\boldsymbol{M}(\boldsymbol{r}, \mathrm{t})/\mu_0$.

Concerning the units, sometimes the so-called *cgs-* (centimeter, gram, second) *system* is used. Though it has only marginal differences in mechanics to the *SI system*, which is based on the units meter, kilogram, second, Ampere, Kelvin, mol and candela, the cgs-system produces non-intuitive units in electrodynamics (like

the electrostatic unit (esu), which contains the square root of mass). For conversion between the systems see [07B1].

2.2 Electromagnetic Waves in Vacuum

In vacuum the following conditions are fulfilled

$$\boldsymbol{P} = 0\,; \quad \boldsymbol{M} = 0\,; \quad \rho = 0\,; \quad \boldsymbol{j} = 0\,. \tag{2.8}$$

With the help of (2.1e,f) this simplifies (2.1c,d) to

$$\nabla \times \boldsymbol{E} = -\mu_0 \dot{\boldsymbol{H}} \quad \text{and} \quad \nabla \times \boldsymbol{H} = \varepsilon_0 \dot{\boldsymbol{E}} \tag{2.9a,b}$$

Applying $\nabla\times$ to (2.9a) and $\partial/\partial t$ to (2.9b) yields

$$\nabla \times (\nabla \times \boldsymbol{E}) = -\mu_0 \nabla \times \dot{\boldsymbol{H}} \quad \text{and} \quad \nabla \times \dot{\boldsymbol{H}} = \varepsilon_0 \ddot{\boldsymbol{E}}\,. \tag{2.10}$$

From (2.10) we find with the help of the properties of the ∇ operator

$$-\mu_0 \varepsilon_0 \ddot{\boldsymbol{E}} = \nabla \times (\nabla \times \boldsymbol{E}) = \nabla(\nabla \cdot \boldsymbol{E}) - \nabla^2 \boldsymbol{E}\,. \tag{2.11}$$

With (2.8) and (2.1a) we see that

$$\nabla \cdot \boldsymbol{E} = 0 \tag{2.12}$$

and (2.11) reduces to the *wave equation in vacuum*, written here for the electric field

$$\nabla^2 \boldsymbol{E} - \mu_0 \varepsilon_0 \ddot{\boldsymbol{E}} = 0\,. \tag{2.13}$$

An analogous wave equation can be obtained for the magnetic field strength.

For monochromatic plane waves (2.17) the general wave equation (2.13) for the electric field can be simplified to the vectorial *Helmholtz equation*

$$(\nabla^2 + \boldsymbol{k}^2)\boldsymbol{E}(\boldsymbol{r}) = 0 \tag{2.14}$$

or to a scalar Helmholtz equation for situations, where the vector character is of no importance.

Solutions of the wave equation (2.13) are all waves of the form

$$\boldsymbol{E}(\boldsymbol{r}, t) = \boldsymbol{E}_0 f(\boldsymbol{k}\boldsymbol{r} - \omega t)\,. \tag{2.15}$$

with $\boldsymbol{E}_0$ being the amplitude and f is an arbitrary function the second derivative of which exists. As can be shown by inserting the ansatz (2.15) into (2.13) the *wave*

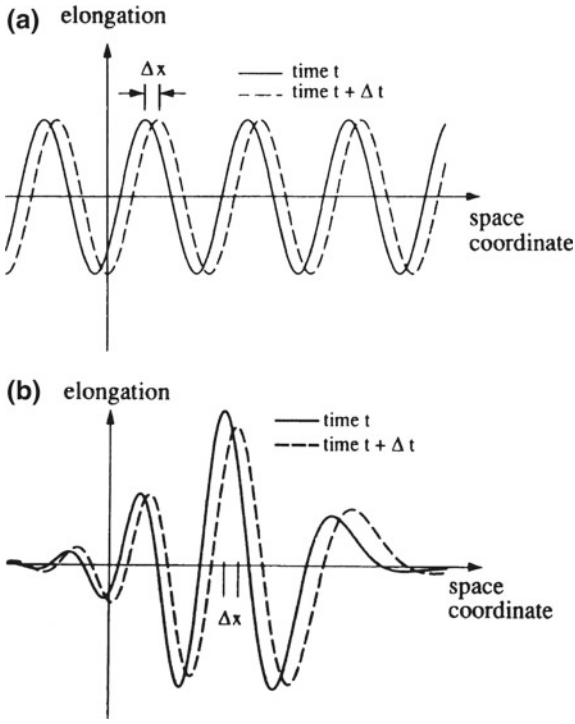


Fig. 2.1 A harmonic wave (**a**) and a wave packet (**b**) shown at two different times t and $t + \Delta t$ to illustrate the concepts of phase and group velocity, respectively

vector $\boldsymbol{k}$ and the *angular frequency* ω obey the relation

$$\frac{\omega}{k} = \left(\frac{1}{\mu_0 \varepsilon_0}\right)^{1/2} = c \quad \text{with} \quad k = |\boldsymbol{k}| = 2\pi/\lambda_{\text{vac}} \,. \tag{2.16}$$

In the following we use for simplicity only the term "frequency" instead of "angular frequency" for $\omega = 2\pi/T$ where T is the *temporal period of the oscillation*. In (2.16), c is the vacuum *speed of light* and λ_{vac} is the *wavelength* in vacuum. From all possible solutions of the form (2.15) we shall concentrate in the following on the most simple ones, namely on *plane harmonic waves*, which can be written as

$$\boldsymbol{E}(\boldsymbol{r}, t) = \boldsymbol{E}_0 \exp[\mathrm{i}(\boldsymbol{k}\boldsymbol{r} - \omega t)]. \tag{2.17}$$

For all waves (not only those in vacuum), the *phase* and *group velocities* v_{ph} and v_{g} are given by

$$v_{\text{ph}} = \frac{\omega}{k}; \quad \boldsymbol{v}_{\text{g}} = \frac{\partial \omega}{\partial k} = \text{grad}_{\boldsymbol{k}} \omega \,, \tag{2.18}$$

where v_{ph} gives the velocity with which a certain phase propagates, (e.g., a maximum of a monochromatic wave) while v_{g} gives the speed of the center of mass of a *wave*

packet with middle frequency ω and covering a small frequency interval $\mathrm{d}\omega$ as shown schematically in Fig. 2.1a,b, respectively. The formulas (2.18) are of general validity. The $\mathrm{grad}_{\boldsymbol{k}}$ on the r.h.s. of (2.18) means a differentiation with respect to $\boldsymbol{k}$ in the sense of $\boldsymbol{\nabla}_k = (\partial/\partial k_x, \partial/\partial k_y, \partial/\partial k_z)$ and has to be used in anisotropic media instead of the more simple expression $\partial\omega/\partial k$. For the special case of electromagnetic radiation in vacuum we find from (2.16) and (2.18)

$$v_{\mathrm{ph}} = v_{\mathrm{g}} = c = (\mu_0 \varepsilon_0)^{-1/2}\,. \tag{2.19}$$

Now we want to see which properties follow from Maxwell's equations for various quantities such as $\boldsymbol{E}_0$ and $\boldsymbol{k}$. Inserting (2.15) or (2.17) into (2.12) gives

$$\boldsymbol{\nabla} \cdot \boldsymbol{E} = \mathrm{i}\boldsymbol{E}_0 \cdot \boldsymbol{k} \exp[\mathrm{i}(\boldsymbol{kr} - \omega t)] = 0. \tag{2.20}$$

This means for the amplitude vector of the electric field that

$$\boldsymbol{E}_0 \perp \boldsymbol{k} \tag{2.21}$$

or, in other words, the electromagnetic wave is *transverse* in $\boldsymbol{E}$.

What can we deduce from Maxwell's equations for the other fields? From (2.9) we get for plane waves

$$\boldsymbol{H} = (\omega\mu_0)^{-1}\boldsymbol{k} \times \boldsymbol{E} = \boldsymbol{H}_0 \exp[\mathrm{i}(\boldsymbol{kr} - \omega t)] \tag{2.22a}$$

with

$$\boldsymbol{H}_0 = (\omega\mu_0)^{-1}\boldsymbol{k} \times \boldsymbol{E}_0. \tag{2.22b}$$

Furthermore we have with (2.1e,f) and (2.8)

$$\boldsymbol{D} = \boldsymbol{D}_0 \exp[\mathrm{i}(\boldsymbol{kr} - \omega t)] = \varepsilon_0 \boldsymbol{E}_0 \exp[\mathrm{i}(\boldsymbol{kr} - \omega t)], \tag{2.22c}$$

$$\boldsymbol{B} = \boldsymbol{B}_0 \exp[\mathrm{i}(\boldsymbol{kr} - \omega t)] = \omega^{-1}\boldsymbol{k} \times \boldsymbol{E}_0 \exp[\mathrm{i}(\boldsymbol{kr} - \omega t)]. \tag{2.22d}$$

The electromagnetic wave is, according to (2.22b), also transverse in $\boldsymbol{B}$ and the electric and magnetic fields are perpendicular to each other, that is, we have in general

$$\boldsymbol{D} \perp \boldsymbol{k} \perp \boldsymbol{B} \perp \boldsymbol{D}. \tag{2.22e}$$

In vacuum and isotropic media one has in addition

$$\boldsymbol{E} \parallel \boldsymbol{D} \quad \text{and} \quad \boldsymbol{H} \parallel \boldsymbol{B}. \tag{2.22f}$$

As we shall see later in the discussion of light propagation in mater, the electric field in (2.20) has to be replaced by the electric displacement field (see (2.46)).

Still, one usually has transverse waves in matter which obey (2.22e) but additionally, longitudinal waves exist under certain conditions.

The *momentum density* $\boldsymbol{\Pi}$ of the electromagnetic field is given by

$$\boldsymbol{\Pi} = \boldsymbol{D} \times \boldsymbol{B}, \quad \boldsymbol{\Pi} \parallel \boldsymbol{k} \tag{2.23}$$

and the *energy flux density* by the *Poynting vector* $\boldsymbol{S}$

$$\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H} \tag{2.24}$$

with $\boldsymbol{S} \parallel \boldsymbol{\Pi}$ in vacuum and isotropic materials. The Poynting vector $\boldsymbol{S}$ is a rapidly oscillating function of space and time. The temporally averaged value $\langle \boldsymbol{S} \rangle$ is usually called the *intensity* I or the *energy flux density*. The intensity is proportional to the amplitude squared for all harmonic waves. For the plane monochromatic waves treated here, we obtain

$$\langle \boldsymbol{S} \rangle = I = \frac{1}{2} |\boldsymbol{E}_0 \times \boldsymbol{H}_0| = \frac{1}{2} c \varepsilon_0 \boldsymbol{E}_0^2 = \frac{1}{2} \frac{c}{\mu_0} \boldsymbol{B}_0^2 = \frac{1}{2} c \mu_0 \boldsymbol{H}_0^2 . \tag{2.25}$$

Equations 2.23 and 2.24 are also valid in matter.

Since a plane electromagnetic wave is a transverse vector wave (2.22e), it can be polarized (see e.g., Sect. 5.1). There are two basis sets, namely either two orthogonally linearly polarized waves or two counter-rotating circularly polarized ones. Both sets of bases can be transformed into each other in the sense that the superposition of a left and right circular polarized wave with equal ω, $\boldsymbol{k}$ and amplitude $\boldsymbol{E}_0$ and with relative phase shifts of 0 or $\pm\pi/2$ results in two orthogonal, linearly polarized waves, while the superposition of two linear polarized waves again with equal ω, $\boldsymbol{k}$ and amplitude $\boldsymbol{E}_0$ and with phase shifts of $\pm\pi/4$ results in a right and left circularly polarized wave.

In a similar way, all types of elliptically polarized waves can be created with different amplitudes of the two partial waves and phase shifts different from $\pm\pi/4$ or $\pm\pi/2$.

For details including the formal description of the polarization state by the Jones vector and the action of optical devices by the Jones matrix for completely polarized light and by the Stokes parameters and the Mueller matrix for partly polarized light see e.g., [07B1, 07M1, 07S1].

2.3 Electromagnetic Waves in Matter

Now we want to elaborate the properties of light in matter. To do so we have in principle to use Maxwell's equations in their general from (2.1). However we will still make in several steps some assumptions which are reasonable for semiconductors: we assume first that there are no macroscopic external space charges i.e. $\rho = 0$. Then a treatment of (2.1) in analogy with (2.9)–(2.13) results in

$$\nabla^2 \boldsymbol{E} - \mu_0 \varepsilon_0 \ddot{\boldsymbol{E}} = \mu_0 \ddot{\boldsymbol{P}} + \mu_0 \dot{\boldsymbol{j}} + \nabla \times \dot{\boldsymbol{M}}\,, \tag{2.26}$$

Actually, there also occurs in the derivation of (2.26) a term $\nabla \cdot (\nabla \boldsymbol{P})$ or $\nabla \cdot (\nabla \boldsymbol{E})$. This term vanishes for transverse waves and is therefore neglected if not mentioned otherwise.

Equation (2.26) is the inhomogeneous analogue of (2.13) telling us that the sources of an electro-magnetic radiation field can be

- a dipole moment $\boldsymbol{p}$ or a polarization $\boldsymbol{P}$ with a non-vanishing second time derivative
- a temporally varying current density
- the curl of a temporally varying magnetization.

Again a similar equation can be obtained for the magnetic field.

We continue now with the application of further simplifications and assume that we have a nonmagnetic material, i.e., that the third term on the r.h.s. of (2.26) vanishes. Actually, all matter has some diamagnetism. But this is a rather small effect of the order of 10^{-6} so it can be neglected for our purposes. Paramagnetic and especially ferromagnetic contributions can be significantly larger for low frequencies. However, even these contributions diminish rapidly for higher frequencies in homogeneous media. Consequently the assumption of a nonmagnetic material is a good approximation over a wide range of the electromagnetic spectrum starting in the IR even for ferromagnetic materials. Furthermore, the most common semiconductors are not ferro-, ferri- or antiferromagnetic and have only a small concentration of paramagnetic centres which may be seen in electron paramagnetic resonance (EPR), but which have negligible influence on the optical properties. The only exceptions are semiconductors which contain a considerable amount of e.g., Mn or Fe ions as does $Zn_{1-y}Mn_ySe$. We refer the reader to, e.g., [88F1, 94D1, 98O1, 04K1] or to Sect. 24.1 for this class of materials.

The *current* term $\boldsymbol{j}$ in (2.1d) deserves some more consideration. The current is driven by the electric field

$$\boldsymbol{j} = \sigma \boldsymbol{E}, \tag{2.27}$$

where σ is the *conductivity*. For intrinsic or weakly doped semiconductors, the carrier density is small and consequently σ is as well. Then the following inequality holds

$$|\boldsymbol{j}| = |\sigma \boldsymbol{E}| \ll |\dot{\boldsymbol{D}}|. \tag{2.28}$$

In the following we will consider this case and neglect $\boldsymbol{j}$ in (2.1d). For heavily doped semiconductors (2.28) is no longer valid and σ will have some influence on the optical properties at least in the infrared (IR). We will come back to this situation in connection with plasmons in Chap. 19.

With these simplifications the basic material equation still left in comparison with the vacuum case is now (2.1e): $\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P}$. If we proceed with this equation again in the manner of (2.9)–(2.13) the result is

$$\nabla^2 \boldsymbol{E} - \mu_0 \varepsilon_0 \ddot{\boldsymbol{E}} = \mu_0 \ddot{\boldsymbol{P}}\,. \tag{2.29}$$

Equation 2.29 states the well-known fact that every dipole moment $\boldsymbol{p}$ and every polarization density $\boldsymbol{P}$ with a non vanishing second derivative in time radiates an electromagnetic wave.

As long as we have no detailed knowledge about the relationships between $\boldsymbol{D}$, $\boldsymbol{E}$ and $\boldsymbol{P}$ we cannot go beyond (2.29). Now we make a very important assumption. We assume a linear relationship between $\boldsymbol{P}$ and $\boldsymbol{E}$:

$$\frac{1}{\varepsilon_0}\boldsymbol{P} = \chi \boldsymbol{E} \tag{2.30a}$$

or

$$\boldsymbol{D} = \varepsilon_0(1+\chi)\boldsymbol{E} = \varepsilon\varepsilon_0\boldsymbol{E} \tag{2.30b}$$

with

$$\varepsilon = \chi + 1. \tag{2.30c}$$

This linear relation is the basic assumption of this first volume of the book on Semiconductor Optics which deals with so-called *linear optics*.

In principle we can also consider (2.30a) as an expansion of $\boldsymbol{P}(\boldsymbol{E})$ in a power series in $\boldsymbol{E}$ which is truncated after the linear term. We will come back to this aspect in the second volume of this book where nonlinear optical properties will be treated in detail.[3] The quantities ε and χ are called the *dielectric function* and the *susceptibility*, respectively. They can be considered as *linear response functions* (see Appendix A or [93S1, 09H1]). Both quantities actually depend on the frequency ω and on the wave vector $\boldsymbol{k}$, and they both have a real and an imaginary part as shown below for ε.

$$\varepsilon = \varepsilon(\omega, \boldsymbol{k}); \quad \chi = \chi(\omega, \boldsymbol{k}) = \varepsilon(\omega, \boldsymbol{k}) - 1, \tag{2.31}$$

$$\varepsilon(\omega, \boldsymbol{k}) = \varepsilon_1(\omega, \boldsymbol{k}) + \mathrm{i}\varepsilon_2(\omega, \boldsymbol{k}). \tag{2.32}$$

The frequency dependence is dominant and will be treated first in Chaps. 5–7. We drop the $\boldsymbol{k}$ dependence for the moment but come back to it in connection with the phenomenon of spatial dispersion in Chap. 7. In the latter we will discuss the properties of ε as a function of frequency and wave vector or as a function of time and space. The value of $\varepsilon(\omega)$ for $\omega \simeq 0$ is usually called the *dielectric constant*.

In general ε and χ are tensors which is easy to comprehend for crystalline materials. For simplicity we shall consider them to be scalar quantities if not stated otherwise. We will discuss the consequences of the tensorial character of the optical functions e.g., in connection with optical anisotropy in Sect. 5.4.

[3] A constant term in this power expansion such as $\boldsymbol{P} = \boldsymbol{P}_0 + \chi E$ would describe a spontaneous polarization of matter which occurs e.g., in pyro- or ferro-electric materials. With arguments similar to the ones given for ferromagnetics we can neglect such phenomena in the discussion of the optical properties of semiconductors.

Using the described assumptions and the linear relations of (2.30) we can transform (2.26) into the *wave equation for the electric field in a dielectric medium*

$$\nabla^2 \boldsymbol{E} - \mu_0 \varepsilon_0 \varepsilon(\omega) \ddot{\boldsymbol{E}} = 0, \tag{2.33a}$$

where we assumed also that $\varepsilon(\omega)$ is spatially constant on a length scale of the order of the wavelength of light. Deviations of this assumption occur e.g., for spatially structured materials like photonic crystals which are treated in the second volume of this textbook.

If magnetic properties are to be included, a corresponding linear approach would lead to

$$\nabla^2 \boldsymbol{E} - \mu_0 \mu(\omega) \varepsilon_0 \varepsilon(\omega) \ddot{\boldsymbol{E}} = 0, \tag{2.33b}$$

where $\mu(\omega)$ is the *magnetic permeability*. As outlined above for most semiconductors in the visible spectral range holds $\mu(\omega) \simeq 1$.

As for (2.13) the solutions of (2.33a) are again all functions of the type

$$\boldsymbol{E} = \boldsymbol{E}_0 f(\boldsymbol{k}\boldsymbol{r} - \omega t), \tag{2.34}$$

or for our present purposes, i.e. again for the case of plane harmonic waves

$$\boldsymbol{E} = \boldsymbol{E}_0 \exp[\mathrm{i}(\boldsymbol{k}\boldsymbol{r} - \omega t)]. \tag{2.35}$$

However, the relationship between $\boldsymbol{k}$ and ω is now significantly different from (2.16). It follows again from inserting the ansatz (2.34) or (2.35) into (2.33a) and now reads:

$$\frac{c^2 \boldsymbol{k}^2}{\omega^2} = \varepsilon(\omega). \tag{2.36}$$

This relation, which relates the macroscopic wave properties $\boldsymbol{k}$ and ω to the characteristic material property $\varepsilon(\omega)$ is called *polariton equation* and will be discussed again in Chap. 7 where we deduce a microscopic model for the light-matter interaction represented by $\varepsilon(\omega)$. The relation (2.36) can also be written in other forms:

$$\boldsymbol{k} = \frac{\omega}{c} \varepsilon^{1/2}(\omega) = \frac{2\pi}{\lambda_{\mathrm{vac}}} \varepsilon^{1/2}(\omega) = \boldsymbol{k}_{\mathrm{vac}}\, \varepsilon^{1/2}(\omega)\,, \tag{2.37}$$

where λ_{vac} and $\boldsymbol{k}_{\mathrm{vac}}$ refer to the vacuum values of these wave quantities.

For the square root of ε we introduce for simplicity a new quantity $\tilde{n}(\omega)$ which we call the *complex index of refraction*

$$\tilde{n}(\omega) = n(\omega) + \mathrm{i}\kappa(\omega) = \varepsilon^{1/2}(\omega). \tag{2.38a}$$

Comparison to (2.32) yields the following relations between the various real and imaginary parts of the optical functions

$$\varepsilon_1(\omega) = n^2(\omega) - \kappa^2(\omega) \tag{2.38b}$$

$$\varepsilon_2(\omega) = 2\, n(\omega)\, \kappa(\omega) \tag{2.38c}$$

The equations (2.16) and (2.36)–(2.38) can be interpreted in the following way. In vacuum an electromagnetic wave propagates with a wave vector $\boldsymbol{k}_{\text{vac}}$ which is real and given by (2.16). In matter, light propagates with a wave vector $\boldsymbol{k}$ which is a complex quantity given by (2.37), or, with the help of (2.38), by

$$k = \frac{\omega}{c}\tilde{n}(\omega) = \frac{\omega}{c}n(\omega) + \mathrm{i}\frac{\omega}{c}\kappa(\omega) = \frac{2\pi}{\lambda_{\text{vac}}}\tilde{n}(\omega) = k_{\text{vac}}\tilde{n}. \tag{2.39}$$

We should notice that k is for complex $\tilde{n}$ not simply $|\boldsymbol{k}|$ since $|\boldsymbol{k}|$ is always a positive, real quantity. Here k means just neglecting the vector character of $\boldsymbol{k}$ but k can still be a real, imaginary or complex quantity according to (2.39). The direction of the real part of $\boldsymbol{k}$, which describes the oscillation of the wave in space, is still parallel to the momentum density $\boldsymbol{D} \times \boldsymbol{B}$ as in (2.23).

Writing down the plane wave explicitly we have:

$$\boldsymbol{E}_0 \exp[\mathrm{i}(\boldsymbol{kr} - \omega t)] = \boldsymbol{E}_0 \exp\left\{\mathrm{i}\left[\frac{\omega}{c}n(\omega)\hat{\boldsymbol{k}}\boldsymbol{r} - \omega t\right]\right\} \exp\left[-\frac{\omega}{c}\kappa(\omega)\hat{\boldsymbol{k}}\boldsymbol{r}\right], \tag{2.40}$$

where $\hat{\boldsymbol{k}}$ is the unit vector in the direction of $\boldsymbol{k}$, i.e., in the direction of propagation of the planes of constant phase. Obviously $n(\omega)\hat{\boldsymbol{k}}$ describes the oscillatory spatial propagation of light in matter; $n(\omega)$ is often called the *refractive index* e.g., in connection with Snell's law of refraction. We find that the *wavelength* λ in a medium is altered with respect to the wavelength λ_{vac} in vacuum to

$$\lambda = \lambda_{\text{vac}} n^{-1}(\omega). \tag{2.41}$$

The second exponential term in (2.40) has a real argument and describes the damping of the wave in the direction of propagation. This effect is usually called *absorption* or, more precisely, *extinction* and $\kappa(\omega)$ is the *extinction coefficient*. We will describe the precise meaning of the two mechanisms in Sect. 5.2. Here we compare (2.40) with the well-known *law of absorption* for the light intensity I of a parallel beam propagating in z-direction

$$I(z) \;=\; I(0)\, \mathrm{e}^{-\alpha z} \tag{2.42a}$$

with I following from (2.25) to be now

$$I = \frac{1}{2}\frac{n(\omega)}{c\mu_0}\boldsymbol{E}_0^2 = \frac{1}{2}\varepsilon_0 c n(\omega)\boldsymbol{E}_0^2 \tag{2.42b}$$

where $\alpha(\omega)$ is usually called the *absorption coefficient*.[4]

Bearing in mind that the intensity is still proportional to the amplitude squared (2.25), a comparison between (2.40) and (2.42) yields

$$\alpha(\omega) = \frac{2\omega}{c}\kappa(\omega) = \frac{4\pi}{\lambda_{\text{vac}}}\kappa(\omega). \tag{2.43}$$

The phase velocity of light in a medium is now given with the use of (2.18) by

$$v_{\text{ph}} = \frac{\omega}{\text{Re}\{k\}} = cn^{-1}(\omega). \tag{2.44}$$

For the group velocity we can get rather complicated dependencies originating from

$$v_{\text{g}} = \frac{\partial\omega}{\partial k}. \tag{2.45}$$

We will return to an illustration of these two velocities later.

We want to discuss now possible solutions of the wave equation in matter (2.33a) which are not found in vacuum. The only solution of (2.12) for light in vacuum is a transverse electromagnetic wave (2.22). This solution exists for light in matter as well. However (2.12) has now with the use of (2.30b) the form

$$\nabla \cdot \boldsymbol{D} = \nabla\varepsilon_0\varepsilon(\omega)\boldsymbol{E} = 0 \tag{2.46}$$

Apart from the above-mentioned transverse solution with $\boldsymbol{E} \perp \boldsymbol{k}$ there is a new solution which does not exist in vacuum ($\varepsilon_{\text{vac}} \equiv 1$), namely $\varepsilon(\omega) = 0$. This means that we can find longitudinal solutions at the frequencies at which $\varepsilon(\omega)$ vanishes. We call these frequencies correspondingly *longitudinal eigenfrequency* ω_{L} and note that for $\varepsilon(\omega_{\text{L}}) = 0$ a field orientation $\boldsymbol{E} \parallel \boldsymbol{k}$ is possible.

Now let us consider the other fields for this *longitudinal wave* in matter. From (2.30) we see immediately that we have for the longitudinal modes $\boldsymbol{D} = 0$ and $\boldsymbol{E} = -\frac{1}{\varepsilon_0}\boldsymbol{P}$. In matter, the Maxwell's equation $\nabla \times \boldsymbol{E} = -\dot{\boldsymbol{B}}$ is still valid. This leads for plane waves in nonmagnetic material to

$$\boldsymbol{H}_0 = (\omega\mu_0)^{-1}\boldsymbol{k} \times \boldsymbol{E}_0. \tag{2.47}$$

[4] In German literature $\alpha(\omega)$ is also known as "Absorptionskonstante" (*absorption constant*) and dimensionless quantities proportional to $\kappa(\omega)$ are called "Absorptionskoeffizient" or "Absorptionsindex" (*absorption coefficient* or *absorption index*). So some care has to be taken regarding the usage of these terms.

For the longitudinal wave it follows from $\boldsymbol{E} \parallel \boldsymbol{k}$ for the magnetic fields that $\boldsymbol{H} = 0$ and $\boldsymbol{B} = \mu_0 \boldsymbol{H} = 0$. These longitudinal waves which we have found in matter are not electromagnetic waves but pure *polarization waves* with $\boldsymbol{E}$ and $\boldsymbol{P}$ opposed to each other and with vanishing fields $\boldsymbol{D}$, $\boldsymbol{B}$ and $\boldsymbol{H}$.

At this point it is appropriate to shortly address the properties of *metamaterials* and *materials with near-zero index*.

On the basis of (2.33b) on actually obtains

$$\tilde{n}(\omega) = \pm(\varepsilon(\omega)\mu(\omega))^{1/2}\,. \tag{2.48}$$

The + sign has to be used for all usual materials, for which $\mu(\omega) > 0$ holds. Recently a new class of artificially designed materials has been invented, the so-called left-handed or metamaterials. They have in some spectral range $\mu(\omega) < 0$ and $\varepsilon(\omega) < 0$. Since group and phase velocity are now pointing in opposite directions, the − sign has to be used, resulting in phenomena like negative refraction, perfect lensing or optical cloaking. For reviews on metamaterials see, e.g., [00P1, 04S1, 17M1]

In near-zero index materials the dielectric function $\epsilon(\omega) \approx 0$ or the permeability $\mu(\omega) \approx 0$ or even both are close to zero. Each condition leads according to (2.48) to a vanishing refractive index. For light waves Maxwell's equations (2.1)c,d, yield $\nabla \times \boldsymbol{E} = \mathrm{i}\omega\mu H \approx 0$ and $\nabla \times \boldsymbol{H} = -\mathrm{i}\omega\epsilon E \approx 0$, i.e., the electric and magnetic field and also wavelength and frequency of the light wave are uncoupled. The phase velocity given by (2.44) diverges leading for monochromatic waves to a spatially static field distribution. The group velocity (see (2.45)) depends on the dispersion of the material. Interesting scenarios occur when v_g is close to zero leading to slow light or when it is close to the light velocity information and energy is transported even for static field distributions. One tries to achieve near-zero index materials by respective tailoring of metamaterials. But also some doped transparent semiconductors like AZO and ITO show near-zero parameters close to the plasma frequency which lies close to the near-infrared communication wavelenghts. For reviews of these materials and their applications see, e.g., [17A1, 17L1].

Until now we have considered the properties of light in the bulk of a medium. The boundary of this medium will need some extra consideration e.g., the interface between vacuum (air) and a semiconductor. This interface is crucial for reflection and refraction of light which we will discuss in detail in Sects. 5.1, and 21.1. Here we only want to state that the boundary conditions allow a surface mode, that is, a wave which propagates along the interface and has field amplitudes which decay exponentially on both sides. These waves are also known as *surface polaritons* for reasons discussed in more detail in Sect. 21.1.5.

Problems

2.1 The intensity of the sunlight falling on the earth is, for normal incidence and before its passage through the atmosphere, about 1.5 kW m^{-2}. Calculate the electric-field strength.

2.2 Pulsed high power lasers can be easily focused to a power density I of 10 GW/cm^2. Calculate the E and B fields. Compare them with the electric field in an H atom at a distance of one Bohr radius, and the magnetic field on the surface of the earth, respectively.

2.3 Show qualitatively the $\boldsymbol{B}$, $\boldsymbol{H}$ and $\boldsymbol{M}$ fields of a homogeneously magnetized, brick-shaped piece of iron and for a hollow sphere with inner radius R_0 and outer radius $R_0 + \Delta R$, which is radially magnetized. Use especially for the second case symmetry considerations together with (2.1).

2.4 Compare the contribution of the electric conductivity of a typical semiconductor to that of the polarisation in (2.26) or (2.28). For which frequencies does the second one dominate?

2.5 Write down the time and space dependence of a spherical wave. Note that the energy flux density varies usually like the amplitude squared. Is it possible to create a spherical vector wave?

2.6 Inspect (with the help of a textbook or a computer program) the electric field of a static electric dipole and the near and far fields of an oscillating electric dipole. Note that in the near field the electric and magnetic fields are not orthogonal.

2.7 Consider or find in a textbook the pattern of the collective motion of the H_2O molecules in a surface water wave. Which effects contribute to the restoring force? Are water waves harmonic waves? What happens at a seashore, where the depth of the water decreases gradually? Is there a net transport of matter? Assume that the particles have an electric charge and move relative to a fixed background of opposite charge. Which charge pattern do you expect close to the surface? Compare with Fig. 7.4b.

References

[65M1] J.C. Maxwell, Philos. Trans. R. Soc. Lond. **155**, 459 (1865)
[88F1] J.K. Furdyna, J. Kossut (eds.), *Diluted Magnetic Semiconductors, Semiconductors and Semimetals*, vol. 25 (Academic, Boston, 1988)
[91B1] I.N. Bronstein, K.A. Semendjajew, *Taschenbuch der Mathematik*, 25th edn. (Teubner, Stuttgart, 1991)
[92S1] H. Stöcker (ed.), *Taschenbuch Mathematischer Formeln und Moderner Verfahren* (Deutsch, Frankfurt, 1992)

[93S1] W. Stößel, *Fourieroptik* (Springer, Berlin, 1993)
[94D1] T. Dietl, Diluted magnetic semiconductors, in *Handbook of Semiconductors*, vol. 38, ed. by T.S. Moss (North Holland, New York, 1994)
[98O1] H. Ohno, Science **281**, 951 (1998)
[00P1] J.B. Pendry, Phys. Rev. Lett. **85**, 3966 (2000)
[04K1] H. Kalt, in *Semiconductor Quantum Structures, Optical Properties*, vol. III/34C2, Landolt-Börnstein, New Series, ed. by C. Klingshirn (Springer, Berlin, 2004)
[04S1] D.R. Smith, J.B. Pendry, M.C.K. Wiltshire, Science **305**, 788 (2004)
[07B1] Y.B. Band, *Light and Matter: Electromagnetism, Optics, Spectroscopy and Lasers* (Wiley, Chichester, 2007)
[07M1] D. Meschede, *Optics, Light and Lasers*, 2nd edn. (Wiley-VCH, Weinheim, 2007)
[07S1] B.E.A. Saleh, M.C. Teich, *Fundamentals of Photonics*, 2nd edn. (Wiley, Hoboken, 2007)
[09H1] H. Haug, S.W. Koch, *Quantum Theory of the Optical and Electrical Properties of Semiconductors*, 5th edn. (World Scientific, Singapore, 2009)
[10C1] K. Cho, *Reconstruction of Macroscopic Maxwell Equations*, Springer Tracts in Modern Physics, vol. 237 (Springer, Berlin, 2010)
[17A1] A. Alù, N. Engheta, J. Opt. (United Kingdom) **19**, 084008 (2017)
[17L1] I. Liberal, N. Engheta, Nat. Photonics **11**, 149 (2017)
[17M1] S. Tretyakov, A. Urbas, N. Zheludev (eds.), Special Issue on the History of Metamaterials. J. Opt. (United Kingdom) **19** (2017)

Chapter 3
Photons

Abstract In this chapter we introduce the concept of photons and shortly describe some basic particle-related properties of light. We will illustrate also the importance of dispersion relations for the understanding of particle properties. It should be noted, that we will discuss here the quantum aspects of light on an elementary level only. We will come back to a comprehensive review of photon physics and the area of quantum optics in the second volume of this book.

3.1 Quantum Aspects of Light

Maxwell's equations are together with the Lorentz force (2.1a)–(2.1g) the basis of the classical theory of light. Using these equations we can deduce properties of light resulting from its wave nature. These properties include e.g., propagation of light, interference or its diffraction at objects like a slit or an optical grating. Well known theoretical approaches to describe these wave phenomena are Huygens principle, Kirchhoff's diffraction theory or Fourier optics [93S1, 07M1, 07S1, 16H1].

The quantum nature of light becomes apparent e.g., in its interaction with matter. The *photoelectric effect*, elaborated experimentally by Heinrich Hertz and Wilhelm Hallwachs in 1887 while performing the first studies on electromagnetic radiation, shows that a light field of frequency ω can exchange energy with matter only in quanta $\hbar\omega$. Such a *quantization* of the light field was first introduced mathematically in 1900 by Max Planck in his formula describing thermal radiation of a black body. Albert Einstein then introduced in 1905 the hypothesis of light quantization in his Nobel Prize- honored explanation of the photoelectric effect. The term *Photon* was made popular in the 1920 ties by Arthur H. Compton, who described the interaction of high-energy light with electrons in terms of particle scattering (Compton effect).

Therefore, a comprehensive theoretical description of light has to be done in terms of quantum mechanics and quantum electrodynamics (QED). The particle nature of light actually is in the focus of topical research activities in the area of *Photonics* or *Quantum Optics*. This research covers a large range from aspects of quantum statistics determining the degree of coherence of light sources, via the preparation and use of single photons in quantum information science or the modification of

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_3

light-mater interactions in cavity-QED. In the following we will shortly introduce the concept of the photon and some of its basic properties. We will return to more detailed discussions e.g., of quantum optics in the second volume of this book.

3.2 Quantization of the Electromagnetic Field

In electrodynamics the electromagnetic fields can be described using their potentials $\boldsymbol{A}$ and ϕ by

$$\boldsymbol{E} = -\nabla\phi - \dot{\boldsymbol{A}}; \quad \boldsymbol{B} = \nabla \times \boldsymbol{A}\,. \tag{3.1}$$

where $\boldsymbol{A}$ is the so-called *vector potential*. Since $\nabla \cdot (\nabla \times \boldsymbol{A}) \equiv 0$ the notation of (3.1) fulfills automatically $\nabla \cdot \boldsymbol{B} = 0$ and reduces the six components of $\boldsymbol{E}$ and $\boldsymbol{B}$ to four.

The vector potential $\boldsymbol{A}$ is not exactly defined by (3.1). A gradient of a scalar field can be added. We can choose the so-called *Coulomb gauge*

$$\nabla \cdot \boldsymbol{A} = 0\,. \tag{3.2}$$

In this case ϕ is the usual *electrostatic potential* obeying the *Poisson equation*:

$$\nabla^2\phi = -\frac{\rho}{\varepsilon_0\varepsilon(\omega)}\,. \tag{3.3}$$

In vacuum we still have $\rho = 0$ and we assume the same for the description of the optical properties of matter.

Now we will shortly sketch the procedure of *second quantization* leading to the photon states and their properties. We use here for simplicity again the plane wave description of light. A detailed treatment of second quantization and its application is beyond the scope of this book so we refer to [00L1, 06F1, 07M2] for further reading.

We start from the classical Hamilton function H which is the total energy of the electromagnetic field using $\boldsymbol{A}$ and ϕ. Then we must find some new, suitable quantities $p_{\boldsymbol{k},s}$ and $q_{\boldsymbol{k},s}$ which are linear in $\boldsymbol{A}$ and which fulfill the canonical equations of motion

$$\frac{\partial H}{\partial q_{\boldsymbol{k},s}} = -\dot{p}_{\boldsymbol{k},s}, \quad \frac{\partial H}{\partial p_{\boldsymbol{k},s}} = -\dot{q}_{\boldsymbol{k},s} \tag{3.4}$$

and are thus canonically conjugate variables. Here $\boldsymbol{k}$ is the wave vector of the plane electromagnetic or $\boldsymbol{A}$-wave and s the two possible transverse polarization directions.

The Hamilton function reads in these variables:

$$H = \frac{1}{2}\sum_{\boldsymbol{k},s}\left(p^2_{\boldsymbol{k},s}\right) + \omega_k^2 q^2_{\boldsymbol{k},s}\ , \tag{3.5}$$

which is the usual form of the Hamiltonian of the *harmonic oscillator*. The quantization condition

$$p_{\boldsymbol{k},s} q_{\boldsymbol{k},s} - q_{\boldsymbol{k},s} p_{\boldsymbol{k},s} = \frac{\hbar}{\mathrm{i}} \tag{3.6}$$

for all $\boldsymbol{k}$ and $s = 1, 2$ gives then the well-known result for the harmonic oscillator: The electromagnetic radiation field has for every $\boldsymbol{k}$ and polarization s discrete energies

$$E_{\boldsymbol{k}} = \left(n_{\boldsymbol{k}} + \frac{1}{2} \right) \hbar \omega_{\boldsymbol{k}} \quad \text{with} \quad n_{\boldsymbol{k}} = 0, 1, 2 \ldots \tag{3.7}$$

It can exchange energy with other systems only in units of $\hbar\omega$. These energy units or quanta are called photons. The term $\hbar\omega/2$ in (3.7) is the zero-point energy of every mode of the electromagnetic field.

The so-called *particle-wave dualism*, that is, the fact that light propagates like a wave showing, e.g., diffraction or interference and interacts with matter via particle-like quanta, can be solved by the simple picture that light is an electromagnetic wave, the amplitude of which can have only discrete values so that the energy in the waves just fulfills (3.7).

From the above introduced, or better, postulated quantities $p_{\boldsymbol{k},s}$ and $q_{\boldsymbol{k},s}$ we can derive by linear combinations operators $a^\dagger_{\boldsymbol{k},s}$ and $a_{\boldsymbol{k},s}$ with the following properties: If $a_{\boldsymbol{k},s}$ acts on a state which contains $n_{\boldsymbol{k},s}$ quanta of momentum $\boldsymbol{k}$ and polarization s it produces a new state with $n_{\boldsymbol{k},s}-1$ quanta. Correspondingly, $a^\dagger_{\boldsymbol{k},s}$ increases $n_{\boldsymbol{k},s}$ by one. We call therefore $a_{\boldsymbol{k},s}$ and $a^\dagger_{\boldsymbol{k},s}$ *annihilation* and *creation operators*, respectively. Since the operators $a_{\boldsymbol{k},s}$ and $a^\dagger_{\boldsymbol{k},s}$ describe *bosons* (see below), their *commutation relation* is

$$\left[a_{\boldsymbol{k},s}, a^\dagger_{\boldsymbol{k},s} \right] = a_{\boldsymbol{k},s} a^\dagger_{\boldsymbol{k},s} - a^\dagger_{\boldsymbol{k},s} a_{\boldsymbol{k},s} = 1 \,. \tag{3.8a}$$

This holds for equal $\boldsymbol{k}$ and s. The *commutator* is zero otherwise.

The operator $a^\dagger_{\boldsymbol{k},s} a_{\boldsymbol{k},s}$ acting on a photon state gives the number of photons $n_{\boldsymbol{k},s}$ times the photon state and is therefore called the *number operator*. Summing over all possible $\boldsymbol{k}$-values and polarizations s gives finally the Hamilton operator

$$H = \sum_{\boldsymbol{k},s} \hbar \omega_{\boldsymbol{k},s} a^\dagger_{\boldsymbol{k},s} a_{\boldsymbol{k},s} . \tag{3.8b}$$

Since a more detailed discussion of quantum electrodynamics is beyond the scope of the first volume of this book, we want to stress here only the important implications of the second quantization: the electromagnetic radiation field in vacuum can be brought into a mathematical form analogous to that of the harmonic oscillator. Quantum mechanics gives for every harmonic oscillator the energetically equidistant terms of (3.7). These are important findings since the harmonic oscillator is one of the most intuitive and also best understood fundamental systems in physics. In the course of this book we will come across many other systems which are treated

analogous to harmonic oscillators. For those readers who are not familiar with the concept of quantization, we strongly recommend the study of a related text book like [08M1].

Here are some more results of the theoretical considerations: A single quantum of the electromagnetic field – i.e., a single photon – carries a *spin* component in the direction of propagation $s_z = \pm\hbar$. The two basic polarizations of single photons are therefore left and right circular (σ^- and σ^+), respectively. A linearly polarized light wave can be considered as a coherent superposition of a left and a right circularly polarized one with equal frequencies, amplitudes and wave vectors $\boldsymbol{k}$. The term coherent means here that two light waves have a fixed phase relation relative to each other. The component of the angular momentum s in the direction of the quantization axis (which is parallel to $\boldsymbol{k}$) is for photons thus as stated already above

$$s_z = s_\parallel = \pm\hbar. \tag{3.9}$$

This means that photons have integer spin and are bosons. The third possibility $s_\parallel = 0$ expected for spin-one particles is forbidden, because longitudinal electromagnetic waves do not exist at least in vacuum.

Photons in thermodynamic equilibrium are described by Bose–Einstein statistics. The occupation probability f_{BE} of a bosonic state with frequency ω is given by

$$f_{\mathrm{BE}} = \left[\exp(\hbar\omega/k_{\mathrm{B}}T) - 1\right]^{-1}, \tag{3.10}$$

where T is the absolute temperature and k_{B} is Boltzmann's constant. The chemical potential μ which actually could appear in the argument of the exponential term in (3.10) is zero in thermal equilibrium, since the number of photons is not conserved. Approaches to describe non-thermal photon fields, e.g., luminescence by a non-vanishing chemical potential and Kirchhoff's law in the sense of a generalized Planck's law are found in [82W1, 95D1, 09H1].

The momentum $\boldsymbol{p}$ of a photon with wave vector $\boldsymbol{k}$ is given, as for all quanta of harmonic waves, by

$$\boldsymbol{p} = \hbar\boldsymbol{k}. \tag{3.11}$$

where $\boldsymbol{k}$ is the real part of the wave vector, which describes as already mentioned the oscillatory, propagating aspect of the plane wave.

To summarize, we can state that photons are bosons with spin $\pm\hbar$, energy $\hbar\omega$ and momentum $\hbar\boldsymbol{k}$ which propagate according to the wave equations.

3.3 Dispersion Relations

Energy ($\hbar\omega$) and (quasi-)momentum ($\hbar\boldsymbol{k}$) are properties of quantum mechanical (quasi-)particles[1] for which conservation laws apply. Further, the quantities ω and $\boldsymbol{k}$

[1] For the concept of quasi-particles and quasi-momenta in crystalline solids see Chap. 11.

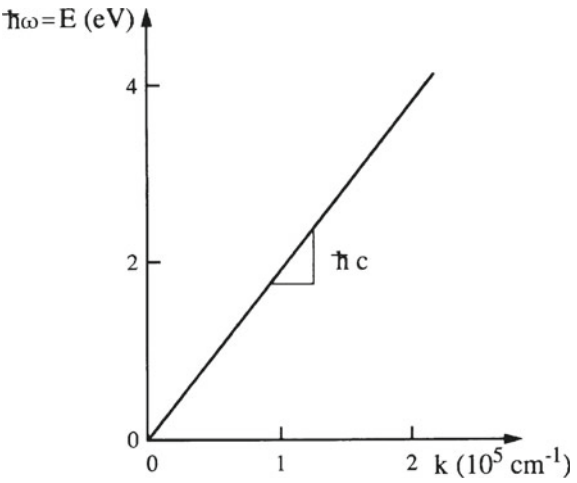


Fig. 3.1 The dispersion relation $E(k)$ of photons in vacuum

describe the periodicity in time and space of the associated wave. Thus it is obvious that dispersion relations given by the dependence of energy E or frequency ω on the wave vector $\boldsymbol{k}$ i.e., the $E(\boldsymbol{k})$ or $\omega(\boldsymbol{k})$, are useful relations to describe the propagation of particles/waves.

For photons in vacuum we find the classical relation given already in (2.16)

$$E = \hbar\omega = \hbar ck. \tag{3.12}$$

The dispersion relation for photons in vacuum is thus a linear function with slope $\hbar c$ as shown in Fig. 3.1. Correspondingly we find again for both, phase and group velocity, with (2.19)

$$v_{\mathrm{ph}} = v_{\mathrm{g}} = c\,. \tag{3.13}$$

Please note that the unit given for the energy on the axis of ordinate in Fig. 3.1 is (eV). We will use this unit throughout this book since it is a more convenient unit for quanta in optical spectroscopy than Joule[2] as can be seen from:

$$1\ \mathrm{eV} = 1.60217733 \times 10^{-19}\ \mathrm{J} \approx 1.6 \times 10^{-19}\ \mathrm{J}\,. \tag{3.14}$$

The energy 1 eV is gained by an electron if it passes, in vacuum, through a potential difference of one Volt. The energy of photons in the visible spectrum extends from ∼1.6 to ∼3.1 eV.

In spectroscopy another measure of energy which is frequently used is the *wave number*. The definition is as follows. One expresses the energy of a (quasi-)particle by the number of wavelengths per cm of a photon in vacuum with the same energy.

$$1\ \mathrm{eV} \mathrel{\widehat{=}} 8065.4\ \mathrm{cm}^{-1} \text{ or } 10^4\ \mathrm{cm}^{-1} \mathrel{\widehat{=}} 1.23986\ \mathrm{eV}\,. \tag{3.15}$$

[2] In the SI system the energy unit Joule is given by $1\ \mathrm{J} = 1\ \mathrm{N\,m} = 1\ \mathrm{kg\,m^2/s^2} = 1\ \mathrm{W\,s} = 1\ \mathrm{V\,A\,s}$.

Unfortunately the term "wave number" is also used for the amount of the real part of the wave vector $k = |\boldsymbol{k}|$ (see Sect. 2.2). Both quantities have the same dimension 1/length, but different definitions which is illustrated by the following example: an optical phonon (see Chaps. 11 and 12) can have for a wave vector zero, i.e., $|\boldsymbol{k}| = 0\,\mathrm{cm}^{-1}$, an energy of e.g., $500\,\mathrm{cm}^{-1}$.

On the axis of abscissae in Fig. 3.1 we plot the amount of the real part of the wave vector which is given by $k = 2\pi/\lambda$, where λ is the wavelength of the corresponding quantum or particle (electron, phonon, photon, etc.). The direction of $\boldsymbol{k}$ is the direction of propagation, i.e., $\boldsymbol{k}$ is normal to the wave front, in the case of light in vacuum or in matter normal to $\boldsymbol{D} \times \boldsymbol{B}$. The amount of the wave vector of light is in the visible in vacuum falling in the range of 10^5 cm^{-1} while the border of the first Brillouin zone (see Chap. 10) defining the maximum meaningful amount of wave vector of a quasi-particle in a periodic lattice is of the order of 10^8 cm^{-1}.

Problems

3.1 Calculate the number of photon modes in the visible part of the spectrum ($\approx$400 nm $\lesssim \lambda_\mathrm{v} \lesssim$ 800 nm) in a box of 1 cm^3.

3.2 Calculate the momentum and energy of a photon with $\lambda_\mathrm{vac} = 500$ nm. At which acceleration voltage has an electron the same momentum?

3.3 Show that the definition $\boldsymbol{v}_\mathrm{g} = \frac{1}{\hbar}\nabla_\mathrm{k} E(\boldsymbol{k})$ leads, for massive and massless particles, directly to the relation (quasi-)momentum $\boldsymbol{p} = \hbar\boldsymbol{k}$.

References

[82W1] P. Würfel, J. Phys. C **15**, 3967 (1982)
[93S1] W. Stößel, *Fourieroptik* (Springer, Berlin, 1993)
[95D1] E. Daub, P. Würfel, Phys. Rev. Lett. **74**, 1020 (1995)
[00L1] R. Loudon, *The Quantum Theory of Light*, 3rd edn. (Oxford University Press, Oxford, 2000)
[06F1] M. Fox, *Quantum Optics: An Introduction*, Oxford Master Series in Physics (Oxford University Press, Oxford, 2006)
[07M1] D. Meschede, *Opics, Ligth and Lasers*, 2nd edn. (Wiley-VCH, Weinheim, 2007)
[07M2] P. Meystre, M. Sargent III, *Elements of Quantum Optics* (Springer, Berlin, 2007)
[07S1] B.E.A. Saleh, M.C. Teich, *Fundamentals of Photonics*, 2nd edn. (Wiley, Hoboken, 2007)
[08M1] D.A.B. Miller, *Quantummechanics for Scientists and Engineers* (Cambridge University Press, Cambridge, 2008)
[09H1] K. Henneberger, Phys. Status Solidi B **246**, 283 (2009)
[16H1] E. Hecht, *Optics*, 5th edn. (Pearson, Upper Saddle River, 2016)

Chapter 4
Density of States and Distribution Functions

Abstract The *density of states* is a rather important property not only for photons but for any (quasi-)particle. Thus we will discuss this quantity in a rather general way which can then also be used throughout this book for the statistics of any type of (quasi-)particles including systems of reduced dimensionality. To describe the probability that the available states are occupied with (quasi-)particles we will introduce respective *distribution functions* for fermionic and bosonic particles.

Quantum-mechanical particles occupy discrete states characterized by quantum numbers. Often these states are spaced very densely, i.e. quasi-continuously, in energy or (quasi-)momentum. In particular in many-particle systems it is useful not to look at the individual states but at the density of states. This quantity enters, e.g., in the calculation of the carrier density in an electronic band or in Fermi's golden rule which allows one to calculate transition probabilities and rates between bands.

If we consider a particle which is described by a wave function $\psi(\boldsymbol{r})$ then the probability w to find it in a small element of space $\mathrm{d}\tau = \mathrm{d}x\mathrm{d}y\mathrm{d}z$ around $\boldsymbol{r}$ is

$$w(\boldsymbol{r})d\tau = \psi^*(\boldsymbol{r})\psi(\boldsymbol{r})\mathrm{d}\tau \tag{4.1a}$$

Since the particle has to be somewhere in the system, $w(\boldsymbol{r})$ has to be *normalized*, that is,

$$\int_{\text{system}} w(\boldsymbol{r})\mathrm{d}\tau = \int_{\text{system}} \psi^*(\boldsymbol{r})\psi(\boldsymbol{r})\mathrm{d}\tau = 1. \tag{4.1b}$$

Here, the functions $\psi(\boldsymbol{r})$ are of the form $\exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r})$. For normalization purposes a factor Ω has to be multiplied

$$\psi(\boldsymbol{r}) = \Omega^{-1/2}\exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r}). \tag{4.2}$$

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_4

The normalization condition (4.1b) results in

$$\Omega^{-1}\int_{\text{system}} \exp(-\mathrm{i}\boldsymbol{kr})\exp(\mathrm{i}\boldsymbol{kr})\mathrm{d}\tau = \Omega^{-1}\int_{\text{system}} \mathrm{d}\tau = \Omega^{-1}V_{\text{system}} = 1, \qquad (4.3)$$

where V_{system} is the volume of the physical system. Consequently the normalization factor Ω is just identical to the volume of the system. To avoid a factor of zero in front of the plane-wave term, one assumes that the system is so big that it contains all physically relevant parts, but that it is not infinite. The simplest choice is a box of length L, or, more precisely speaking, a cube in three dimensions, a square in two, and an interval in one. This procedure is known as *normalization in a box*. Consequently we have

$$V_{\text{system}} = L^d \quad \text{with} \quad d = \text{dimensionality of the system} \qquad (4.4)$$

and

$$\Omega^{1/2} = L^{d/2} \quad \text{for} \quad d = 3, 2, 1. \qquad (4.5)$$

The wave vectors which can exist in such a box are limited by the boundary conditions.

If we assume that we have an infinitely high potential barrier around the box, then the wavefunction must have nodes at the walls (Fig. 4.1a). Consequently the

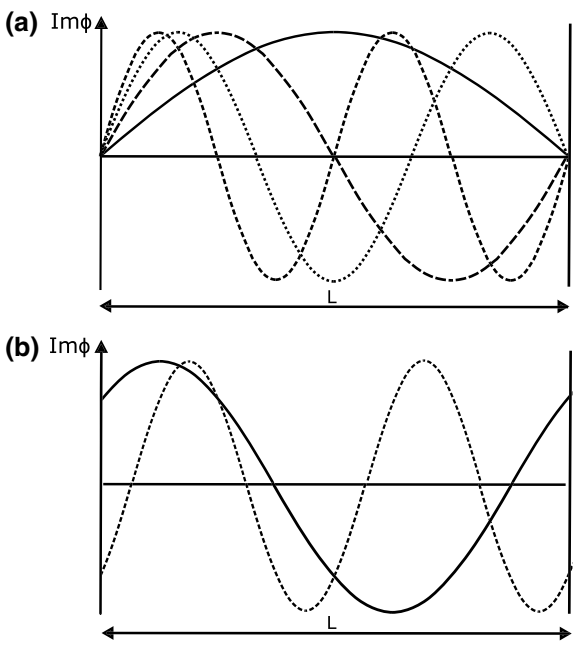


Fig. 4.1 Plane waves which have nodes at the boundaries (**a**) or which obey periodic boundary conditions (**b**)

components k_i of $\boldsymbol{k}$ must fulfill

$$k_i = n_i \frac{\pi}{L}; \quad n_i = 1, 2, 3, \ldots; \quad i = 1, \ldots, d, \tag{4.6}$$

where the index i runs over all dimensions.

Such a wave is a standing wave, i.e., a coherent superposition of two waves with $\boldsymbol{k}$ and $-\boldsymbol{k}$ and equal amplitudes. In the following we must consider therefore only positive values of $\boldsymbol{k}$. The various modes are distributed equally spaced over the k_i-axes with a spacing Δk_i given by

$$\Delta k_i = \frac{\pi}{L}. \tag{4.7}$$

In other words, every state (or mode) needs a volume V_k in $\boldsymbol{k}$-space given by

$$V_k = \left(\frac{\pi}{L}\right)^d. \tag{4.8}$$

Another approach is to impose *periodic boundary conditions*. Then the plane wave should have equal amplitude and slope on opposite sides of the cube according to Fig. 4.1b. In this case one can fill the infinite space by adding boxes in all dimensions and one finds:

$$k_i = n_i' \frac{2\pi}{L}; \quad n_i' = \pm 1, \pm 2, \pm 3 \ldots \tag{4.9}$$

This means for Δk_i

$$\Delta k_i' = \frac{2\pi}{L}. \tag{4.10}$$

In contrast to the case of standing waves, we now have to consider both positive and negative values of n_i. This procedure results finally in the same density of states.

As a consequence we find that plane waves have in Cartesian coordinates in $\boldsymbol{k}$-space a constant density on all axes.

This result can also be derived qualitatively from the uncertainty principle. If a particle is confined to a length L in direction i its momentum has an uncertainty $\Delta p_i = \Delta \hbar k_i \geq \hbar / L$. Consequently two states, which should be distinguishable or "different", must have k_i values, which differ by roughly $1/L$.

Often one wants to know the number of states in a shell between k and $k + \mathrm{d}k$ independent of the direction of $\boldsymbol{k}$. This question can be answered by introducing polar coordinates in $\boldsymbol{k}$-space. The differential volume $\mathrm{d}V_k$ of a shell of thickness $\mathrm{d}k$ in a d-dimensional $\boldsymbol{k}$-space is given by

$$\begin{aligned} \mathrm{d}V_k &= 2\mathrm{d}k && \text{for } d = 1, \\ \mathrm{d}V_k &= 2\pi k \mathrm{d}k && \text{for } d = 2, \\ \mathrm{d}V_k &= 4\pi k^2 \mathrm{d}k && \text{for } d = 3. \end{aligned} \tag{4.11}$$

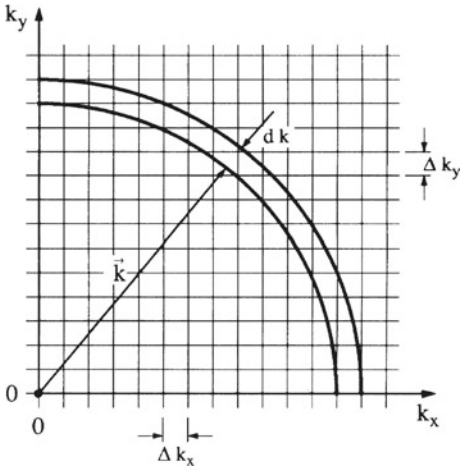


Fig. 4.2 A two-dimensional $\boldsymbol{k}$-space in which the states are equally spaced on the k_x- and k_y-axes to illustrate the derivation of (4.12) for $d = 2$

Depending on the boundary condition we have to take into account only positive (4.6), or positive and negative (4.9), values of $\boldsymbol{k}$ or n_i with corresponding modifications of the prefactors in (4.11).

The number $\hat{D}(\boldsymbol{k})$ of states in $\boldsymbol{k}$-space found between k and $k + \mathrm{d}k$ in polar coordinates is given by dividing $\mathrm{d}V_k$ by the volume for one state and by multiplying by g_s. The quantity g_s considers degeneracies such as the spin degeneracy. For photons we have $g_\mathrm{s} = 2$ according to the σ^+ and σ^- polarizations (see above). The results are

$$\begin{aligned} \hat{D}(k)\mathrm{d}k &= g_\mathrm{s}\frac{L}{\pi}\mathrm{d}k \qquad && \text{for } d = 1, \\ \hat{D}(k)\mathrm{d}k &= g_\mathrm{s}\frac{L^2}{2\pi}k\mathrm{d}k \qquad && \text{for } d = 2, \\ \hat{D}(k)\mathrm{d}k &= g_\mathrm{s}\frac{L^3}{2\pi^2}k^2\mathrm{d}k \qquad && \text{for } d = 3. \end{aligned} \tag{4.12}$$

The derivation of this result is illustrated for $d = 2$ in Fig. 4.2.

If we neglect constant prefactors and divide by $\mathrm{d}k$ we find

$$\hat{D}(k) \propto g_\mathrm{s} L^d k^{d-1}, \quad d = 1, 2, 3 \ldots . \tag{4.13}$$

If we consider not the number of states in the box of volume L^d but the density of states $D(k)$ per unit of space (e.g., per cm^3 or m^3) the term L^d in (4.13) disappears yielding

$$D(k) \propto g_\mathrm{s} k^{d-1} \ . \tag{4.14}$$

This result is expected since the density of states per unit volume must be independent of the size of the box which we have in mind provided the box is sufficiently large.

The concept of periodic boundary conditions yields the same result.

We want to stress here that we assumed only plane waves but did not make any specific assumptions about which type of particles are represented by these plane waves, i.e., photons, electrons etc. Therefore this result is valid for all particles described by plane waves.

Often it is necessary to know the density of states as a function of energy $D(E)\mathrm{d}E$. This quantity gives the number of states in the energy interval from E to $E + \mathrm{d}E$. To calculate this quantity we need the specific dispersion relation $E(\boldsymbol{k})$ and its inverse $\boldsymbol{k}(E)$ as seen from the identity:

$$D(E)\mathrm{d}E = D[k(E)]\frac{\mathrm{d}k}{\mathrm{d}E} \cdot \mathrm{d}E = D[k(E)]\frac{1}{|\nabla_k E(\boldsymbol{k})|}\mathrm{d}E. \tag{4.15}$$

The term on the right-hand side of (4.15) gives the generalized expression which is also valid for anisotropic cases.

In particular for photons in vacuum we have with (3.12)

$$k = \frac{E}{\hbar c} = \frac{\omega}{c}; \quad \frac{\mathrm{d}k}{\mathrm{d}E} = \frac{1}{\hbar c} \tag{4.16}$$

Inserting this result in (4.15), for the case $d = 3$ we find

$$D(E)\mathrm{d}E = \frac{E^2}{\pi^2(\hbar c)^3}\mathrm{d}E \tag{4.17a}$$

or

$$D(\omega)\mathrm{d}\omega \propto \omega^2 \mathrm{d}\omega \ . \tag{4.17b}$$

For massive particles, i.e. particles the dispersion relation of which can be described by an (effective) mass m according to

$$E(\boldsymbol{k}) = \frac{\hbar^2 \boldsymbol{k}^2}{2m} \tag{4.18}$$

we obtain with (4.15) for a d-dimensional space

$$D(E)\mathrm{d}E \propto g_\mathrm{s} E^{\frac{d}{2}-1}\mathrm{d}E \ . \tag{4.19}$$

This formula includes the well-known square-root dependence of the density of states on energy for massive particles in three dimensions.

We emphasize again that the states in a d-dimensional $\boldsymbol{k}$-space are discrete and equidistant in all directions for all particles, which can be described by a plane wave. The density of states as a function of energy depends on the individual dispersion relation. For further reading on the density of states for various (quasi-)particles we refer to [09I1, 14H1].

The density of states describes how many states per unit volume are available for the respective particles. Next we want to describe the probability that these states are actually occupied by an ensemble of particles. We restrict ourselves in the following to thermodynamic equilibrium. The respective distribution function f and the type of statistics to be used depends on the particle properties. There are three types of statistics which can be considered:
– for classical, distinguishable particles, *Boltzmann statistics* apply

$$f_{\mathrm{B}} = \exp[-(E-\mu)/k_{\mathrm{B}}T] \tag{4.20a}$$

– for bosons, i.e., indistinguishable particles with integer spin, photons being an example, one must use the *Bose–Einstein statistics*

$$f_{\mathrm{BE}} = \{\exp[(E-\mu)/k_{\mathrm{B}}T] - 1\}^{-1} \tag{4.20b}$$

– fermions, or indistinguishable particles with half-integer spin, e.g., electrons, obey the *Fermi–Dirac statistics*

$$f_{\mathrm{FD}} = \{\exp[-(E-\mu)/k_{\mathrm{B}}T] + 1\}^{-1}. \tag{4.20c}$$

Here k_{B} is the *Boltzmann constant* and μ the *chemical potential* which gives the average energy necessary to add one more particle to the system. For fermions μ is also known as the *Fermi energy* E_{F}.

The probability to find a particle in the interval from E to $E + \mathrm{d}E$ is then given by the product of the density of states $D(E)$ and the occupation probability f

$$D(E)f(E,T,\mu)\mathrm{d}E \tag{4.21}$$

In Fig. 4.3 we plot f_{B}, f_{BE} and f_{FD} as a function of $(E-\mu)/k_{\mathrm{B}}T$. The Boltzmann distribution function shows the well-known exponential dependence. The Fermi–Dirac distribution never exceeds one in accord with Pauli's exclusion principle. The

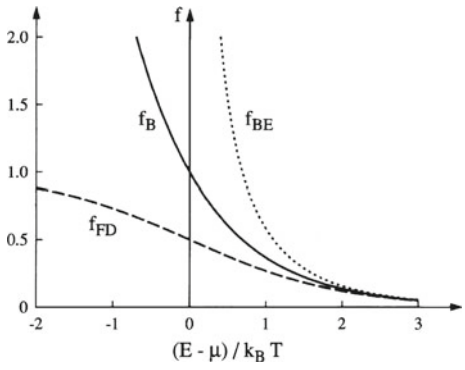


Fig. 4.3 The Boltzmann, Bose–Einstein and Fermi–Dirac distribution functions versus $(E-\mu)/k_{\mathrm{B}}T$

Bose–Einstein distribution has a singularity for $E = \mu$. This allows Bose–Einstein condensation, or in other words, a macroscopic population of a single state, if μ resides in a region with a finite density of states. In this case the species with energies $E = \mu$ and those with $E > \mu$ must be considered separately.

Furthermore it is obvious from Fig. 4.3 that f_{BE} and f_{FD} converge to f_{B} for $(E - \mu)/kT > 1$. So, the classical Boltzmann statistics can be used as a high-energy approximation for Bosons and Fermions. In Sects. 11.2 or 16.3 we will discuss the limits of this approximation and when the use of Fermi–Dirac or Bose–Einstein statistics is obligatory.

The chemical potential μ is zero in thermodynamic equilibrium for bosonic quanta the number of which is not conserved, for e.g., photons or phonons. We will discuss this topic in Chap. 11. If the number N or density n of particles in a system is known, as is the case for electrons at non-relativistic energies, then μ is well defined by:

$$\int D(E) f(E, \mu, T) \mathrm{d}E = n, \tag{4.22}$$

which says that the density of particles is equal to the integral over the product of the density of states and the probability that a state is occupied.

As an example, we apply now the above statements to photons in a three-dimensional box in thermodynamic equilibrium. With (4.18) and (4.20b) we obtain:

$$N(\omega)\mathrm{d}\omega = D(\omega) f_{\mathrm{BE}}(\omega, T)\mathrm{d}\omega \propto \omega^2 \left[\exp\left(\frac{\hbar\omega}{k_{\mathrm{B}}T}\right) - 1\right]^{-1} \mathrm{d}\omega\,. \tag{4.23}$$

For the energy content $U(T)$ of the radiation field this results in

$$U(T) = \hbar \int_0^\infty \omega N(\omega)\mathrm{d}\omega \propto \int_0^\infty \omega^3 \left[\exp\left(\frac{\hbar\omega}{k_{\mathrm{B}}T}\right) - 1\right]^{-1} \mathrm{d}\omega\,. \tag{4.24}$$

Equation 4.24 is nothing else than *Planck's law* of *black-body radiation*. By substituting x for the variable $\hbar\omega/k_{\mathrm{B}}T$ in (4.24) we immediately find the Stefan–Boltzmann T^4 law

$$U(T) \propto T^4 \int_0^\infty x^3 (\exp x - 1)^{-1} \mathrm{d}x = T^4 A\,, \tag{4.25}$$

where A is a constant.

Problems

4.1. Check whether the maximum of $N(\omega)$ in (4.23) shifts in proportion to T (Wien's law), originally formulated as $\lambda_{\max} \propto T^{-1}$.

4.2. How does the density of states as a function of energy vary for a linear dispersion relation (like photons) in 3, 2 and 1 dimensional systems?

References

[09I1] H. Ibach, H. Lüth, *Solid-State Physics*, 4th edn. (Springer, Berlin, 2009)
[14H1] S. Hunklinger, *Festkörperphysik*, 4th edn. (De Gruyter Oldenbourg, Oldenbourg, 2014)

Chapter 5
Macroscopic Optical Properties of Solids

Abstract In this chapter we describe macroscopic optical properties resulting from the interaction of light with solid matter. The solid is considered here as a homogeneous medium described by the complex dielectric function $\varepsilon(\omega)$ or by the complex index of refraction $\tilde{n}(\omega)$. We concentrate especially on the reflection and transmission of light at the plane interface between two media. As an especially simple case we investigate the boundary of matter and vacuum. We deduce the laws of reflection and refraction from boundary conditions for the electric and magnetic fields. The behavior of wave vectors when crossing an interface follows from momentum conservation. Fresnel's formulae are discussed to describe amplitudes, phases and intensities of the respective light waves. It follows the description of extinction of light by absorption and scattering. Since semiconductors often occur as plane-parallel slabs we deduce the properties of Fabry–Perot cavities. The chapter concludes with a discussion of optical anisotropy including birefringence, dichroism and optical activity.

5.1 Reflection and Refraction at an Interface

Light is reflected and refracted when crossing an interface between two media. Starting from the boundary conditions for electric and magnetic fields at an interface we will derive here the laws of reflection and refraction describing the conservation of frequency and change of momentum $\hbar\boldsymbol{k}$ as well as Fresnel's formulae which give the changes of the field amplitude and phase. We will also consider alternative ways to deduce the rules for reflection and refraction.

5.1.1 Boundary Conditions

We start with the macroscopic description of light crossing an interface between two media. The Fig. 5.1 shows wave vectors and field amplitudes in the vicinity of such an interface for two linear polarizations. Note that $\boldsymbol{k}$, $\boldsymbol{E}$ and $\boldsymbol{H}$ form a right-handed system (in contrast: in artificial metamaterials they form a left-handed one, see Sect. 2.3). In Fig. 5.1a the electric field $\boldsymbol{E}_\mathrm{i}$ of the incident beam is polarized parallel to the plane

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_5

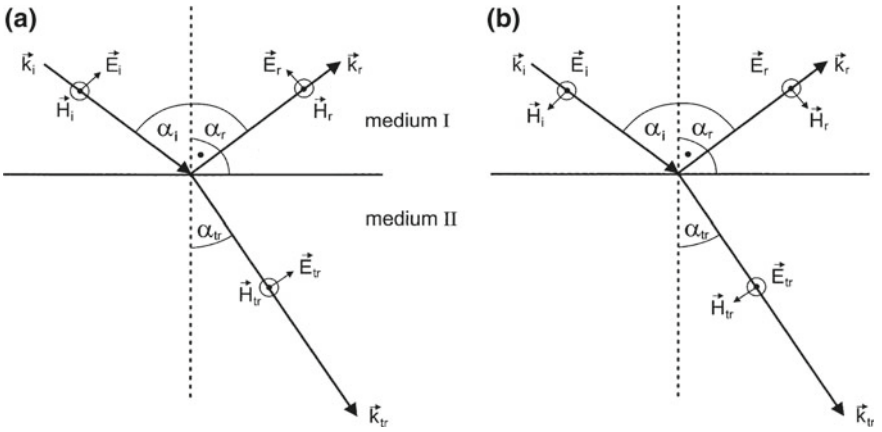


Fig. 5.1 The $\boldsymbol{E}$ and $\boldsymbol{H}$ fields and the wave vectors $\boldsymbol{k}$ for incident, transmitted and reflected beams at an interface between two isotropic media for two different, orthogonal, linear polarizations (**a**) and (**b**) respectively

of incidence, which is defined by the wave vector of the incident light $\boldsymbol{k}_\mathrm{i}$ and the normal to the plane interface $\boldsymbol{e}_\mathrm{n}$. As we will see later from the *boundary conditions*, the wave vectors and the electric fields of transmitted and reflected beams (indices tr and r, respectively) are in the same plane; the magnetic fields are according to (2.22) perpendicular to it. In Fig. 5.1b we have just the opposite situation for $\boldsymbol{E}$ and $\boldsymbol{H}$.

One often assumes for simplicity that the medium I is vacuum (or air), i.e., $\varepsilon_\mathrm{I}(\omega) = \tilde{n}_\mathrm{I}(\omega) \equiv 1$. We do not use this approximation here but we still assume that media I and II are isotropic. This means $\varepsilon_\mathrm{I}(\omega)$ and $\varepsilon_\mathrm{II}(\omega)$ are scalar functions and $\tilde{n}(\omega)$ does not depend on the polarization of the light. Phenomena which appear if we drop this assumption are dealt with later. Furthermore, we assume that there is only one reflected and one transmitted beam. This assumption seems trivial and it is indeed for reflection. In transmission there may be more than one propagating beam, as we shall discuss in Sect. 21.1. What we first want to know are the dependencies of the angles α_r and α_tr on α_i i.e., the laws of reflection and refraction. We then want to know the *coefficients of reflection* r and *transmission* t of the interface between media I and II. We can define these coefficients for the field amplitudes E_0 by

$$r_\parallel = \frac{E_{0\mathrm{r}}}{E_{0\mathrm{i}}}; \quad t_\parallel = \frac{E_{0\mathrm{tr}}}{E_{0\mathrm{i}}} \tag{5.1a}$$

for the configuration of Fig. 5.1a and by

$$r_\perp = \frac{E_{0\mathrm{r}}}{E_{0\mathrm{i}}}; \quad t_\perp = \frac{E_{0\mathrm{tr}}}{E_{0\mathrm{i}}} \tag{5.1b}$$

for the configuration of Fig. 5.1b.

However, what is usually measured is the *reflectivity* R and *transmittivity* T of an interface for the intensities. We have

$$R_{\perp,\parallel} = \mid r_{\perp,\parallel} \mid^2 \tag{5.2}$$

because incident and reflected beams propagate in the same medium. For the definition of the transmittivity we need the considerations of the beam cross-sections which we will discuss later (see Fig. 5.7).

In order to deduce mathematical relationships to calculate the introduced quantities we need boundary conditions which must be fulfilled at the interface. These boundary conditions can be deduced from Maxwell's equations. To do so, we apply two general laws of vector analysis which are known as *Gauss' theorem* and *Stokes' theorem*, respectively (see also (2.7)). They read for a given vector field $\boldsymbol{A}$

$$\int_{\text{volume}} \nabla \cdot \boldsymbol{A} \mathrm{d}\tau = \oint_{\text{surface}} \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{f} \tag{5.3a}$$

and

$$\int_{\text{surface}} (\nabla \times \boldsymbol{A}) \cdot \mathrm{d}\boldsymbol{f} = \oint_{\text{line}} \boldsymbol{A} \cdot \mathrm{d}\boldsymbol{s} \,. \tag{5.3b}$$

Starting from (2.1a) and with the help of (5.3a), we obtain

$$\int_{\text{volume}} \mathrm{div}\boldsymbol{D} \mathrm{d}\tau = \oint_{\text{surface}} \boldsymbol{D} \cdot \mathrm{d}\boldsymbol{f} = \int_{\text{volume}} \rho \mathrm{d}\tau \,. \tag{5.4}$$

We have chosen the integration volume in the form of a tiny (differentially small) cylinder which contains the interface and has its top and bottom in the media I and II, respectively (Fig. 5.2). Furthermore it is assumed that the ratio of the height to the radius of this cylinder is also infinitesimally small, so that the contribution to the whole integral from the lateral surface of the cylinder is negligible compared to the contributions from its top and bottom surfaces. Then the middle and right-hand-side terms of (5.4) yield

$$(\boldsymbol{D}_{\mathrm{I}} - \boldsymbol{D}_{\mathrm{II}}) \cdot \mathrm{d}\boldsymbol{f} = (D_{\mathrm{n,I}} - D_{\mathrm{n,II}}) \mathrm{d}f = \rho_{\mathrm{s}} \mathrm{d}f \,, \tag{5.5}$$

where the index n means the normal component and ρ_{s} a surface charge density. The contribution of a volume charge density ρ_{v} goes to zero with decreasing height of the cylinder as compared to a possible contribution of a two-dimensional surface charge density ρ_{s}.

Since we have assumed that there are no free charges ρ at all (Sect. 2.3) and consequently no surface charges, the right-hand side of (5.5) vanishes and we find as a boundary condition that the normal component of $\boldsymbol{D}$ is continuous across the interface:

$$D_{\mathrm{n}}^{\mathrm{I}} = D_{\mathrm{n}}^{\mathrm{II}} \,. \tag{5.6a}$$

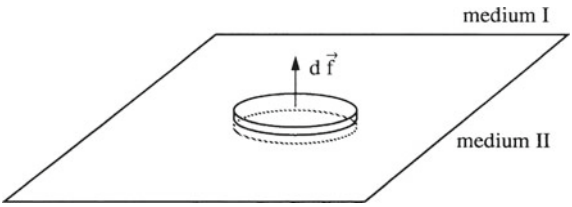


Fig. 5.2 Schematic drawing of the infinitesimally small cylinder used to deduce with (5.4) the boundary condition (5.6)

Starting from (2.1b) we find in the same way

$$B_{\mathrm{n}}^{\mathrm{I}} = B_{\mathrm{n}}^{\mathrm{II}} \,. \tag{5.6b}$$

Using (2.1c, d) and (5.3b) we get in an analogous way requirements for the tangential components of $\boldsymbol{E}$ and $\boldsymbol{H}$.

$$E_{\mathrm{t}}^{\mathrm{I}} = E_{\mathrm{t}}^{\mathrm{II}} \,. \tag{5.7a}$$

$$H_{\mathrm{t}}^{\mathrm{I}} = H_{\mathrm{t}}^{\mathrm{II}} \,. \tag{5.7b}$$

Equations (5.6) and (5.7) represent the boundary conditions for electric and magnetic fields. Actually only two of them are independent, the other ones follow directly with the linear approaches (2.30). We are now able to calculate for a given incident beam the properties of the reflected and refracted ones. In order to do so the boundary conditions must be applied to a specific problem. For the configuration of Fig. 5.1 this reads for the incident, reflected and transmitted electric fields

$$\boldsymbol{e}_{\mathrm{n}} \times \boldsymbol{E}_{\mathrm{i}} - \boldsymbol{e}_{\mathrm{n}} \times \boldsymbol{E}_{\mathrm{r}} = \boldsymbol{e}_{\mathrm{n}} \times \boldsymbol{E}_{\mathrm{tr}} \,, \tag{5.8}$$

since the vector products of the various $\boldsymbol{E}_{\mathrm{s}}$ with the unit vector $\boldsymbol{e}_{\mathrm{n}}$ normal to the interface generate just the tangential components. Using the sets of (5.6) and (5.7) allows one to calculate the properties of the reflected and transmitted beams. This procedure involves some basically simple but lengthy algebra and does not give further insight into the physics. In accordance with the concept of this book, we consequently skip these calculations which can be found in the literature, (see e.g. [76P1, 99B1, 09R1, 16H1] and references therein) but rather present the results giving some cross-links to other physical approaches to obtain them.

5.1.2 The Laws of Reflection and Refraction

The first, not too surprising, result from the above-mentioned procedure is

$$\omega_{\mathrm{i}} = \omega_{\mathrm{r}} = \omega_{\mathrm{tr}} \,. \tag{5.9}$$

This means the light in all three beams has the same frequency. This becomes clear from classical physics, as we shall see in Chap. 7 since atoms perform forced oscillations with frequency ω_i under the influence of the incident field and can therefore radiate, according to the linear approach (2.30), only at this frequency. The relation (5.9) is also intelligible from the point of view of quantum mechanics, bearing in mind the law of energy conservation and the fact that a single photon has energy $\hbar\omega$ and can be either reflected or transmitted.

The next results are the laws of reflection and Snell's law of refraction. The *law of reflection* states

$$\alpha_\mathrm{i} = \alpha_\mathrm{r} \tag{5.10a}$$

and

$$\boldsymbol{k}_\mathrm{i}, \boldsymbol{k}_\mathrm{r} \text{ and } \boldsymbol{e}_\mathrm{n} \text{ are in one plane,} \tag{5.10b}$$

namely in the above-introduced plane of incidence. *Snell's law* reads

$$\frac{\sin\,\alpha_\mathrm{i}}{\sin\,\alpha_\mathrm{tr}} = \frac{n_\mathrm{II}}{n_\mathrm{I}} \tag{5.11a}$$

and

$$\boldsymbol{k}_\mathrm{i}, \boldsymbol{k}_\mathrm{tr} \text{ and } \boldsymbol{e}_\mathrm{n} \text{ are in one plane.} \tag{5.11b}$$

It is often overlooked that the result of all wave-vectors being in the same plane ((5.10b) and (5.11b)) is integral part of each of the above two laws.

In Fig. 5.1 we have illustrated the situation for $n_\mathrm{I} < n_\mathrm{II}$, i.e., the refraction from an optically thinner into an optically thicker medium. In the opposite case, one reaches a critical angle $\alpha_\mathrm{i}^\mathrm{c}$ for which $\alpha_\mathrm{tr} = 90°$ is given by the condition

$$\alpha_\mathrm{i}^\mathrm{c} = \arcsin\frac{n_\mathrm{II}}{n_\mathrm{I}}\,. \tag{5.12}$$

For $\alpha_\mathrm{i} \geq \alpha_\mathrm{i}^\mathrm{c}$ there is a totally reflected beam but no longer a transmitted one. However, the boundary conditions (5.6) and (5.7) require finite field amplitudes in medium II. Inspection of the boundary conditions shows that a so-called *evanescent wave* exists in medium II which propagates parallel to the surface. Its field-amplitudes decay exponentially in the direction normal to the interface over a distance of a few wavelengths, as shown schematically in Fig. 5.3a. The reflected wave has under these conditions the same intensity as the incident one. Correspondingly the phenomenon is known as *total (internal) reflection*.

If medium II has only a thickness of the order of a wavelength and is then covered by material I again, then the evanescent wave couples into this medium giving rise to a propagating, transmitted wave (Fig. 5.3b). Consequently the intensity of the reflected wave decreases. This phenomenon is called *attenuated*, or *frustrated, total*

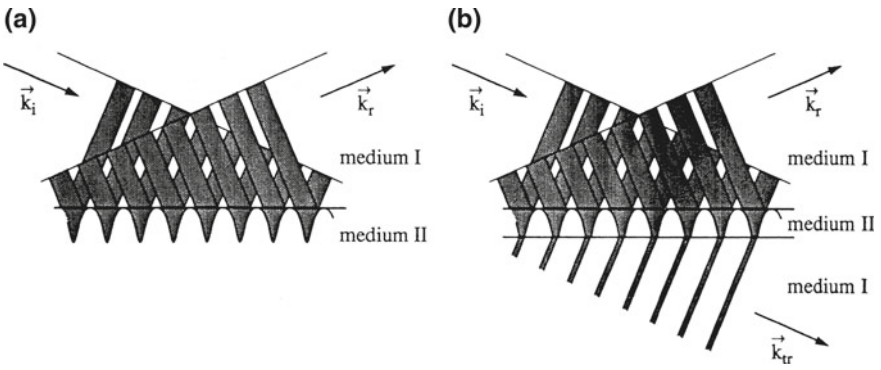


Fig. 5.3 The evanescent wave in the case of total internal reflection (**a**) and the arrangement for the optical tunneling effect or attenuated total reflection (**b**). The widths of the grey bars are intended to symbolize the intensity of the respective waves

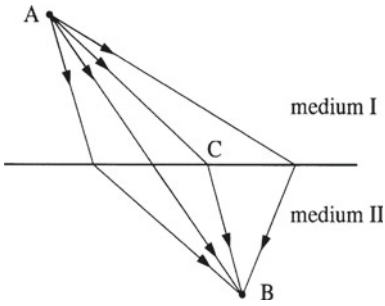


Fig. 5.4 Various possible paths of a light beam travelling from A to B

reflection (ATR) or the *optical tunnel effect* in analogy to the quantum-mechanical tunnel effect [76P1].

The laws of reflection and refraction ((5.10) and (5.11)) can also be deduced from the *principle of Maupertuis* or *Fermat*, which says that for geometrical optics the optical path length, i.e., the product of the geometrical path length and the refractive index n between two points A and B is an extremum, generally a minimum. This is shown schematically for the case of refraction in Fig. 5.4. From all in principle possible ways to travel from A to B, the light propagation is along the one for which it takes the least time, i.e., for which (5.13) holds.

$$\delta \int n \mathrm{d}s = \delta(n_{\mathrm{I}} \overline{AC} + n_{\mathrm{II}} \overline{CB}) = 0\,, \tag{5.13}$$

For the chosen light path the variation δ of the optical path length vanishes. Equation (5.11) can be deduced from (5.13).

5.1.3 Conservation Laws at an Interface

A third way of deriving equations (5.9)–(5.11) relies on the *conservation laws* of some physical quantities. Since these conservation laws are also essential for many other aspects of semiconductor optics we will give here compact introduction to them.

We start with a basic principle of physics formulated Emmy Noether. *Noehter's theorem* says in simple words: *A conservation law follows from every invariance of the Hamilton operator H with respect to a transformation*. We are not going to prove this statement. But we will review here some well known applications and consequences:

- If H is invariant against infinitesimal translations in time $\mathrm{d}t$, i.e., if H does not depend explicitly on time, then the total energy E of the system described by H is conserved

$$H(t) = H(t + \mathrm{d}t) \longrightarrow E = \text{const.} \tag{5.14a}$$

- If H is invariant against an infinitesimal translation along an axis x, then the x-component of the total momentum $\boldsymbol{p}$ is conserved

$$H(x) = H(x + \mathrm{d}x) \longrightarrow p_x = \text{const.} \tag{5.14b}$$

- If H is invariant against an infinitesimal rotation $\mathrm{d}\phi$, e.g., around an axis z, that is, $\mathbf{d}\phi = (0, 0, \mathrm{d}\phi)$, then the z-component of the angular moment $\boldsymbol{L}$ is conserved

$$H(\boldsymbol{r}) = H(\boldsymbol{r} - \boldsymbol{r} \times \mathbf{d}\phi) \longrightarrow L_z = \text{const.} \tag{5.14c}$$

  The axis along which (5.14c) is valid is called the quantization axis.

For the problem of light reflection and refraction at an interface, (5.14a) is strictly valid. Energy conservation results in the conservation of the light frequency ω for reflected and refracted beams (see (5.9)). Invariance of the Hamiltonian (i.e. the physical properties of the system) against infinitesimal translation (5.14b) is fulfilled only parallel to the interface. Correspondingly only the momentum parallel to the interface is conserved. We learned from (3.11) that the momentum of photons (and of all other free particles) is $\hbar\boldsymbol{k}$. Consequently, at the interface the conservation laws must be fulfilled leading to the following properties of the wave-vectors:

$$k_{\mathrm{i}\parallel} = k_{\mathrm{r}\parallel}\,, \tag{5.15a}$$

$$k_{\mathrm{i}\parallel} = k_{\mathrm{tr}\parallel}\,. \tag{5.15b}$$

Since incident and reflected beams propagate in the same medium, the lengths of the wave vectors and thus the wavelengths are equal, too.

$$|\boldsymbol{k}_{\mathrm{i}}| = |\boldsymbol{k}_{\mathrm{r}}|\,, \tag{5.16}$$

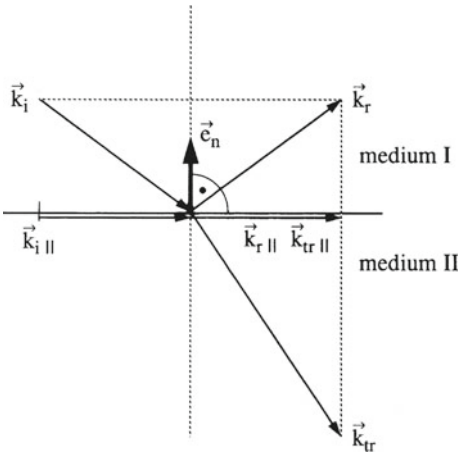


Fig. 5.5 The momenta $\hbar\boldsymbol{k}$ of incident, reflected and refracted beams at an interface

The only possible solution for (5.15a) and (5.16) is then obviously the law of reflection (5.10).

For the relation of the incident $\boldsymbol{k}_\mathrm{i}$ and transmitted $\boldsymbol{k}_\mathrm{tr}$ we have in addition to (5.15b)

$$k_j = k_\mathrm{vac} n_j = \frac{\omega}{c} n_j \,; \quad j = \mathrm{I}, \mathrm{II}. \tag{5.17}$$

The simultaneous solution of (5.15b) and of (5.17) results in (5.11a). Also, since the modulus of the wave-vector changes across an interface, the wavelength of the light changes.

When we describe damping or extinction of the light in a medium by a complex wave vector (2.39) and (2.40) then the above conservation laws apply to the real, i.e., oscillatory part of $\boldsymbol{k}$. For clarity, the situation is depicted (again) in Fig. 5.5.

Finally, the conservation law (5.14c) still holds for a quantization axis perpendicular to the interface.

5.1.4 *Fresnel's Formulae*

So far we have considered what happens to the frequency and wave-vector of light impinging on an interface. We now want to discuss the respective field amplitudes and intensities in the individual beams. From the boundary conditions (5.6), and (5.7) we can deduce the transmission t and the reflection r of an interface for the simplifying assumption that both media are transparent. This means that the imaginary part κ of the complex index of refraction $\tilde{n}$ is negligible, i.e., much smaller than the real part n: $|\kappa| \ll |n| \backsimeq 1$ (condition of weak absorption).

The resulting equations are known as *Fresnel's formulae* for the regime of weak absorption and for non-magnetic materials:

$$r_\perp = \frac{n_\mathrm{I}\cos\alpha_\mathrm{i} - n_\mathrm{II}\cos\alpha_\mathrm{tr}}{n_\mathrm{I}\cos\alpha_\mathrm{i} + n_\mathrm{II}\cos\alpha_\mathrm{tr}} = -\frac{\sin(\alpha_\mathrm{i} - \alpha_\mathrm{tr})}{\sin(\alpha_\mathrm{i} + \alpha_\mathrm{tr})}\,, \tag{5.18a}$$

$$r_\parallel = \frac{-n_\mathrm{II}\cos\alpha_\mathrm{i} + n_\mathrm{I}\cos\alpha_\mathrm{tr}}{n_\mathrm{I}\cos\alpha_\mathrm{tr} + n_\mathrm{II}\cos\alpha_\mathrm{i}} = -\frac{\tan(\alpha_\mathrm{i} - \alpha_\mathrm{tr})}{\tan(\alpha_\mathrm{i} + \alpha_\mathrm{tr})}\,, \tag{5.18b}$$

$$t_\perp = \frac{2n_\mathrm{I}\cos\alpha_\mathrm{i}}{n_\mathrm{I}\cos\alpha_\mathrm{i} + n_\mathrm{II}\cos\alpha_\mathrm{tr}} = \frac{2\sin\alpha_\mathrm{tr}\cos\alpha_\mathrm{i}}{\sin(\alpha_\mathrm{i} + \alpha_\mathrm{tr})}\,, \tag{5.18c}$$

$$t_\parallel = \frac{2n_\mathrm{I}\cos\alpha_\mathrm{i}}{n_\mathrm{I}\cos\alpha_\mathrm{tr} + n_\mathrm{II}\cos\alpha_\mathrm{i}} = \frac{2\sin\alpha_\mathrm{tr}\cos\alpha_\mathrm{i}}{\sin(\alpha_\mathrm{i} + \alpha_\mathrm{tr})\cos(\alpha_\mathrm{i} - \alpha_\mathrm{tr})}\,. \tag{5.18d}$$

The relation between α_i and α_tr according to Snell's law (5.11a) is used to progress from one set of formula to the other. The signs in (5.18) depend on the way in which we defined the field amplitudes in Fig. 5.1. However it is obvious from (5.18a) to (5.18b), that there is for normal incidence ($\alpha_\mathrm{i} = 0°$) a phase shift of 0 or of π between the incident and reflected field for reflection at the optically thinner or thicker medium, respectively, i.e.,

$$r_{\parallel,\perp}(\alpha_\mathrm{i} = 0) = \frac{n_\mathrm{I} - n_\mathrm{II}}{n_\mathrm{I} + n_\mathrm{II}};\ t_{\parallel,\perp}(\alpha_\mathrm{i} = 0) = \frac{2n_\mathrm{I}}{n_\mathrm{I} + n_\mathrm{II}}\,. \tag{5.18e}$$

Actually, $r_\perp$ is negative for all incidence angles α_i while $r_\parallel$ changes from negative to positive values for large always angles α_i. These signs are not relevant as along as one measures only intensities, but they do matter when calculating light interference in layered structures such as Bragg stacks or (anti-)reflection coatings.

We display in Fig. 5.6 formulae (5.18) graphically and the phase shift between the various reflected components, assuming that the incident ones are in phase. Furthermore we show the results for strong absorption not covered by (5.18).

The experimentally accessible quantities are R and T which can be calculated from (2.25) to (5.2). We discuss first the reflectivity R. For the orientation $R_\perp$, the reflectivity increases monotonically with α_i. The limiting values for $\alpha_\mathrm{i} = 0°$ and $\alpha_\mathrm{i} = 90°$ are given in (5.19).

For weak absorption ($n \gg \kappa$)

$$R(\alpha_\mathrm{i}{=}0) = \left(\frac{n_\mathrm{II} - n_\mathrm{I}}{n_\mathrm{II} + n_\mathrm{I}}\right)^2 \tag{5.19a}$$

For $n_\mathrm{I} = 1$ we find for weak absorption

$$R(\alpha_\mathrm{i}{=}0) = \left(\frac{n_\mathrm{II} - 1}{n_\mathrm{II} + 1}\right)^2; \tag{5.19b}$$

and for strong absorption (here the refractive index of medium II has to betreated as a complex quantity [16H1]) and $n_\mathrm{I} = 1$

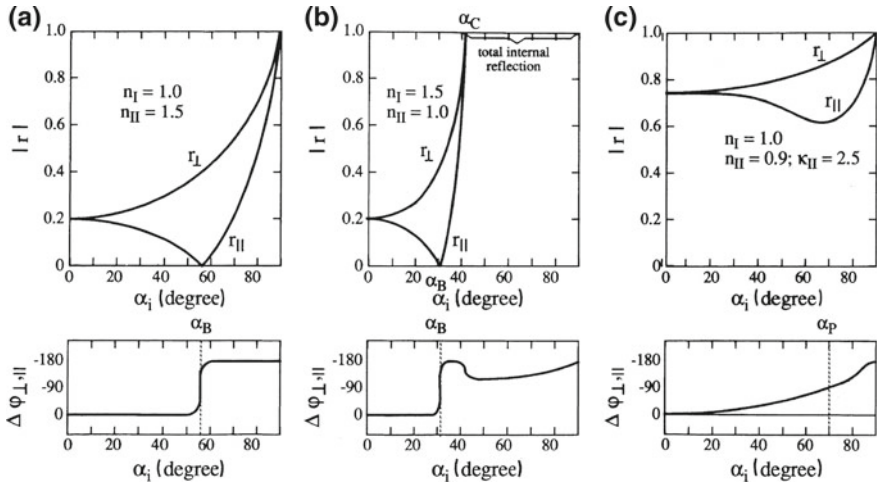


Fig. 5.6 The modulus of the reflection coefficient r for the (electric) field amplitudes according to (5.18) and the phase difference between the reflected amplitudes for the orientations $r_\perp$ and $r_\parallel$, and reflection at an optically thicker (**a**) and thinner (**b**) medium and for reflection at a strongly absorbing medium (**c**) ([76P1], adapted by permission from: Springer Nature, Optik und Atomphysik by R. W. Pohl ©1976)

$$R(\alpha_i{=}0) = \frac{(n_{II} - 1)^2 + \kappa_{II}^2}{(n_{II} + 1)^2 + \kappa_{II}^2}, \tag{5.19c}$$

while for grazing incidence

$$R(\alpha_i{=}90^\circ) = 1 \tag{5.19d}$$

in all cases.

$R_\perp$ increases monotonically with increasing α_i in all cases. In contrast $R_\parallel$ goes through a minimum at a certain angle α_B with

$$R_\parallel(\alpha{=}\alpha_B) = 0 \tag{5.20}$$

for the case of weak absorption. The angle α_B is known as *Brewster's angle* or the polarization angle. For $\alpha = \alpha_B$ only the component polarized perpendicularly to the plane of incidence is reflected. So this angle can be used to polarize light, if unpolarized light is directed to the interface. Note that the transmitted beam is not strictly polarized, but has only some preference for the orientation parallel to the plane of incidence.

From (5.18b) we can deduce the condition for which $r_\parallel = 0$, namely

$$n_{II}\cos\alpha_i = n_I\cos\alpha_{tr} \text{ or } \tan(\alpha_i{+}\alpha_{tr}) = \infty. \tag{5.21}$$

Equation (5.21) has, apart from the trivial solution (no interface $\rightarrow n_{\mathrm{I}} = n_{\mathrm{II}}$; $\alpha_{\mathrm{i}} = \alpha_{\mathrm{tr}}$), the solution

$$\alpha_{\mathrm{i}} + \alpha_{\mathrm{tr}} = 90^{\circ} \ , \tag{5.22}$$

i.e., the reflected and refracted beams propagate perpendicularly to each other. This fact can be easily understood. As we shall see in Chap. 7 the reflected beam is radiated from the forced oscillations of the atoms close to the surface which forms the optically thicker medium (assuming for the moment that medium I is vacuum). Since dipoles do not radiate in the direction of their axis, and since the polarization in the medium is perpendicular to $\boldsymbol{k}_{\mathrm{tr}}$ (transverse wave) we find directly (5.22).

For the case $n_{\mathrm{II}} < n_{\mathrm{I}}$ we also find the critical angle α_{c} for *total internal reflection* in Fig. 5.6b which we mentioned already earlier.

If we send light on the interface polarized differently than $E_{\parallel}$ or $E_{\perp}$ to the plane of incidence, we can decompose it always into two components with the above orientations, we determine their reflected or transmitted amplitudes with (5.18) or from Fig. 5.6 and superpose them again, taking into account the relative phase shifts given in Fig. 5.6. In the general case of a phase shift different from 0° or 180° the reflected light will be elliptically polarized for a linearly polarized incident beam. In experimental investigations of the optical properties of semiconductors, one tries to avoid this additional complication, usually by choosing the simplest geometries. A prominent exception is the experimental method called *ellipsometry* which actually evaluates the polarization changes to determine optical functions over a large spectral range. We will describe this method in Fig. Ell.1.

For strong absorption, R does not reach zero for any polarization (Fig. 5.6c), and starts already for $\alpha_{\mathrm{i}} = 0$ rather close to one. This leads to a statement which may seem contradictory in itself at first glance: strongly absorbing materials absorb only a small fraction of the incident light. The reason for this behavior is simply the fact that the bigger fraction is reflected. The smaller fraction which actually enters the medium is absorbed however over a short distance. In a weakly absorbing medium, the major portion of the light is transmitted through the surface and may be completely absorbed if the medium is thick enough. Indian ink is in the sense of (5.19) a weakly absorbing medium, metals are strongly absorbing over wide spectral ranges and have R close to unity.

With increasing α_{i}, $R_{\perp}$ increases monotonically while $R_{\parallel}$ goes through a shallow minimum as can be inferred from Fig. 5.6c. The so-called *principle angle of incidence* α_{p} is defined by some authors as the α_{i} for which the slopes of the curves $R_{\parallel}(\alpha_{\mathrm{i}})$ and $R_{\perp}(\alpha_{\mathrm{i}})$ are equal. Other authors prefer to use the minimum of $R_{\perp}$ as definition of α_{p}. The difference is marginal. The phase shift between the two components is just $\frac{\pi}{2}$ for α_{p}.

A consequence of the smooth variation of the phase shift with α_{i} is the fact that linearly polarized light impinging on a metallic mirror is usually elliptically polarized after reflection except for the simple orientations $E_{\parallel}$, and $E_{\perp}$. This fact should be remembered when building an optical setup in the lab.

To conclude this subsection, we shall shortly consider transmission through a single interface. For a lossless interface, energy conservation requires for the incident,

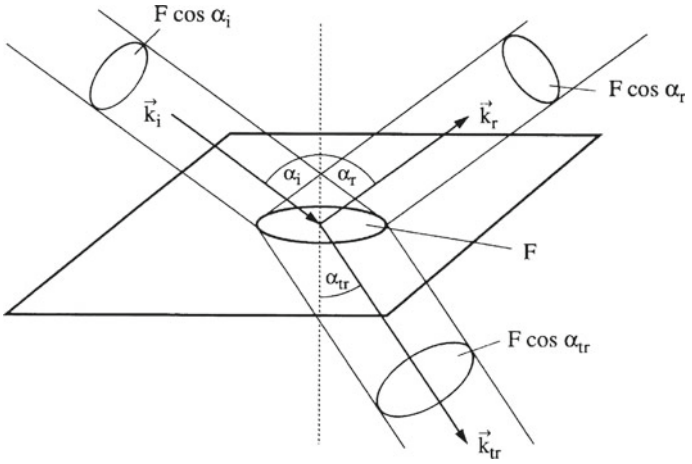


Fig. 5.7 The relation of the cross sections of incident, transmitted and reflected beams at a plane interface between two media

reflected and transmitted powers P of the light

$$P_i = P_r + P_{tr}, \tag{5.23a}$$

where the power is defined as energy per unit of time. Despite (5.23a), $|\ r\ |^2$ and $|\ t\ |^2$ do not add up to unity, since these quantities give information about the reflected and transmitted light intensities. This quantity gives, as stated already earlier, the energy flux density, i.e., the energy per unit of time and of area. Since the cross-sections of the incident and reflected beams are equal, but different from the transmitted one for $\alpha_i \neq 0$ as shown in Fig. 5.7, a corresponding correction factor has to be added to $T = |t|^2$ to fulfill (5.23a):

$$T_{\parallel,\perp} = \frac{P_{tr}}{P_i} = \frac{I_{tr}}{I_i}\frac{\cos\alpha_{tr}}{\cos\alpha_i} = |t_{\parallel,\perp}|^2 \frac{\cos\alpha_{tr}}{\cos\alpha_i}\frac{n_{II}}{n_I}. \tag{5.23b}$$

The last factor on the r.h.s of (5.23b) results from the different light velocities in the media I and II. For more details about transmission and reflection at a plane interface see e.g. [76P1, 09R1, 99B1, 16H1].

5.2 Extinction, Absorption and Scattering of Light

Until now we have considered mainly what happens in the immediate surroundings of a plane interface between two media. Now we will consider the propagation of a beam within a medium, continuing the discussion in Sect. 2.3 in connection with (2.40).

If light propagates in a medium other than vacuum, its intensity or field amplitude decreases in most cases with increasing distance, even if we (un-physically) assume a strictly parallel beam and neglect all diffraction losses connected with its finite diameter. In reality, both assumptions above can be fulfilled only to a good approximation but never in a strict, mathematical sense. The decay is usually exponential with increasing distance d (see (2.40)–(2.43)) and referred to as *Lambert–Beer's law*:

$$I = I_0 \mathrm{e}^{-\alpha(\omega)d} \quad . \tag{5.24}$$

If energy is pumped into a suitable material and in a suitable way $\alpha(\omega)$ may even become negative for a certain range of frequencies and consequently light is amplified. We call these materials *active, laser* or *inverted materials* in contrast to *passive materials* with $\alpha \geq 0$.

The attenuation of light according to (5.24) is called *extinction*. It comprises two groups of phenomena:

$$\alpha_{\text{extinction}}(\omega) = \alpha_{\text{absorption}}(\omega) + \alpha_{\text{scattering}}(\omega) \tag{5.25}$$

Extinction is the more comprehensive term. It enters in the damping γ of the polariton propagation described in Chap. 8 or in the phase relaxation time T_2 of coherent excitations (discussed in more detail in the second volume of this book) via $\gamma = 2\hbar {T_2}^{-1}$ and contains two contributions.

Absorption is the transformation of the energy of the light field into other forms of energy like heat, chemical energy or electromagnetic radiation which is not coherent and generally also frequency shifted with respect to the incident beam. This latter phenomenon is usually called *(photo-)luminescence*.

The other contribution to extinction is attenuation by (coherent) scattering of light. The unshifted component is called *Rayleigh scattering* and requires some disorder in the medium. The frequency-shifted (coherent) parts are known as *Raman-* or *Brillouin scattering*. See Chap. 12 for these effects or [00W1]. If the scattering particles do not show absorption in the visible and have typical sizes large compared to the wavelength of the light λ, the material usually looks white. Examples are the powder of ZnO (just to start with a semiconductor), ground sugar and salt, clouds, snow, the foam of beer, milky quartz or the bark of a birch tree. The reason is that the scattering of light at the interfaces by reflection and refraction is roughly wavelength independent and thus the same for all colors.

If the particles are small compared to λ, the scattering at these particles becomes wavelength dependent. Often one finds an ω^4 law

$$\frac{I_{\text{scatter}}}{I_{\text{incident}}} \propto \omega^4 \tag{5.26}$$

This relation follows from a combination of (2.29) and (2.25). It explains that the sunlight propagating through clear atmosphere preferentially looses the high frequencies, i.e., short wavelength or blue parts of its spectrum by scattering from N_2,

O_2 and other molecules of the atmosphere. Consequently the sun itself looks yellow to red depending on the thickness of air through which the sunlight has to travel and the sky appears blue from the scattered light. The above mentioned disorder necessary for this type of Rayleigh scattering comes from density fluctuations of the particles constituting the air within the coherence volume of (sun-)light. For more details of light scattering, including particles with sizes comparable to the wavelength (the so-called *Mie scattering*) see e.g. [75H1, 99B1, 16H1].

If the scattering or absorbing particles are diluted and do not interact with each other, one finds proportionality between their concentration n_p and α

$$\alpha(\omega) = n_p \alpha_s(\omega), \tag{5.27}$$

where α_s is the specific extinction constant. Equation (5.27) is also known as *Beer's law*. Combined with the finding of J. Lambert that the attenuation of light in a medium is proportional to the path length one can deduce the exponential dependence of the light intensity on the distance called *Lambert–Beer's law* (5.24).

Though there is evidently a rather clear definition of the terms "extinction" and "absorption", one uses often in "every day" language in the lab and also in many books including this one the word absorption instead of extinction, sometimes for convenience, and sometimes because it is not always clear which group of phenomena is responsible for the attenuation of a light beam along its path through matter.

5.3 Transmission Through a Slab of Matter and Fabry–Perot Modes

We discuss now in connection with Figs. 5.8 and 5.9 the transmission and reflection of a *plane-parallel slab* of matter of geometrical thickness d with ideal, lossless surfaces. The surrounding material I is air or vacuum ($n_I = 1$, $\kappa_I = 0$). The total transmission $\hat{T}$ or reflection $\hat{R}$ does not only depend on material II and on the angle α_i but also on the properties of the incident light field, e.g., on its polarization and on its coherence length l_c, i.e., the distance over which there is a fixed phase-relation. We can discuss here only some limiting cases. For further reading see textbooks on Optics and Photonics [76P1, 99B1, 07M1, 07S1, 09R1, 16H1].

For strong absorption in the sense discussed in connection with (5.18) it is easily possible to detect the beam reflected from the front surface. The transmitted beam is strongly attenuated for $d \gg \lambda \simeq 0.5\ \mu m$ in the visible part of the spectrum, i.e., for an optical density $\alpha(\omega)d \gg 1$. Consequently $\hat{R}$ reduces to R given for normal incidence by (5.19c). T is difficult to measure since sufficiently thin samples are often not easily available and sometimes have optical properties different from the bulk material. Similar statements for $\hat{R}$ are true for $\alpha(\omega)\mathrm{d} \gg 1$ even if $\alpha(\omega)$ is small. This can be expressed by

$$\hat{R}(\omega) = R(\omega) \ \text{ for } \ \alpha(\omega)d \gg 1\ . \tag{5.28}$$

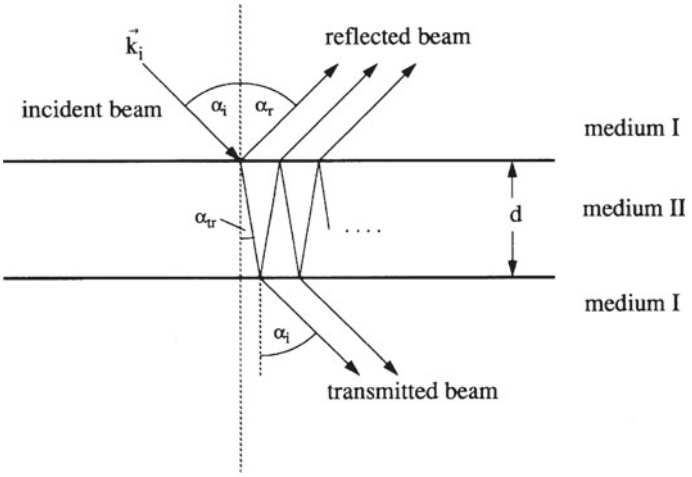


Fig. 5.8 Multiple reflection of an incident light beam in a plane-parallel slab of matter

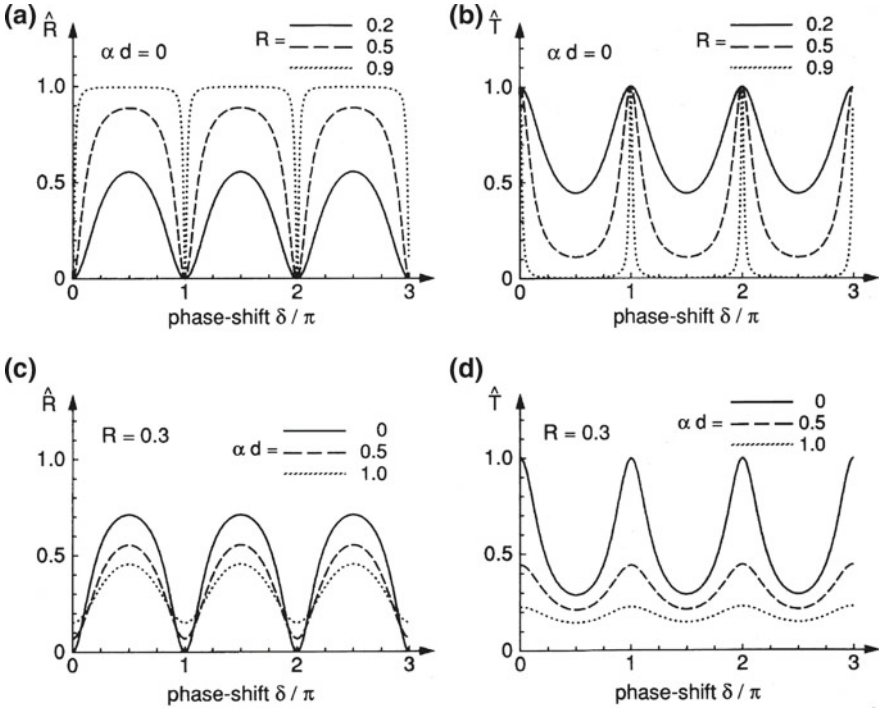


Fig. 5.9 Transmission $\hat{T}$ (**b**, **d**) through and reflection $\hat{R}$ (**a**, **c**) from a Fabry–Perot resonator as a function of the phase shift δ for various values of the reflectivity R of a single surface or the optical density αd of the medium in the resonator

The most convenient regime in which to measure $\alpha(\omega)$ is $1 \leqslant \alpha(\omega)d \leqslant 5$. The reflection has to be taken into account only once at the front and rear surfaces since multiply reflected beams are very weak due to the absorption. We find with Fig. 5.8

$$\hat{T}(\omega) = \left[1 - R^{\mathrm{I}\to\mathrm{II}}(\omega, \alpha_{\mathrm{i}})\right] \exp\left[-\alpha(\omega)\, d \cos^{-1}\alpha_{\mathrm{tr}}\right] \times \left[1 - R^{\mathrm{II}\to\mathrm{I}}(\omega, \alpha_{\mathrm{tr}})\right] \ . \tag{5.29a}$$

This simplifies for normal incidence to (5.29b) bearing in mind that $R^{\mathrm{I}\to\mathrm{II}} = R^{\mathrm{II}\to\mathrm{I}} = R(\omega)$ for weakly absorbing material (see (5.19a)).

$$\hat{T}(\omega) \simeq [1 - R(\omega)]^2 \ \mathrm{e}^{-\alpha(\omega)d} \ . \tag{5.29b}$$

For the conditions of (5.28) we can write

$$\hat{R}(\omega) \simeq R(\omega) + [1 - R(\omega)]^2 \, R(\omega) \, \mathrm{e}^{-2\alpha(\omega)d} \simeq R(\omega) \ . \tag{5.29c}$$

For materials with an optical density $\alpha(\omega)d \leq 1$ the properties of the plane-parallel slab are determined by multiple reflection. The behavior depends strongly on the relation of the optical pathlength and the coherence length l_{c} of the light beam. For $d \gg l_{\mathrm{c}}^{-1}$ (short coherence length) we have to add intensities resulting for normal incidence in all cases discussed by (5.29b)–(5.32) in the following expression:

$$\hat{T}(\omega) \simeq \frac{[1 - R(\omega)]^2 \mathrm{e}^{-\alpha(\omega)d}}{1 - R^2(\omega)\mathrm{e}^{-2\alpha(\omega)d}} \simeq [1 - R(\omega)]^2 \, \mathrm{e}^{-\alpha(\omega)d} \ \text{ for } \ \alpha d \lesssim 1 \, . \tag{5.29d}$$

For long coherence lengths, the field amplitudes interfere with appropriate phases and the parallel-sided slab forms a device called *etalon* or *Fabry–Perot resonator* (abbreviated FP resonator).

There are two limiting cases: in one case all partial waves reflected at the two surfaces interfere constructively in the resonator. This condition is fulfilled if an integer number m of half waves fits in the resonator, expressed mathematically as

$$\lambda_m = \frac{d\, 2n(\omega)}{m} \ \text{ or } \ \omega_m = \frac{m\pi c}{d\, n(\omega_m)} \ \text{ or } \ k_m = m\frac{\pi}{d} \ \text{ with } \ m = 1, 2, 3 \ldots . \tag{5.30}$$

In this case we have a large field amplitude in the resonator which may surpass even the amplitude of the incident beam, a total transmission $\hat{T}$ close to unity and correspondingly a weak total reflection $\hat{R}$ (even if R is close to unity!). In the opposite case of mainly destructive interference of the partial waves in the resonator, we find just the opposite situation, that is, $\hat{R} \leq 1$; $\hat{T} \ll 1$.

The general formula for the FP resonator reads approximately [85C1]

$$\hat{T} = \frac{A}{1 + F \sin^2\delta} \ , \quad \hat{R} = \frac{B + F \sin^2\delta}{1 + F \sin^2\delta} \tag{5.31a}$$

with

$$F = \frac{4R_\alpha}{(1-R_\alpha)^2} ,$$
$$A = \frac{\mathrm{e}^{-\alpha d}(1-R_\mathrm{F})(1-R_\mathrm{B})}{(1-R_\alpha)^2} , \tag{5.31b}$$
$$B = \frac{R_\mathrm{F}(1-R_\alpha/R_\mathrm{F})}{(1-R_\alpha)^2} ,$$

R_F: reflectivity of front surface of the Fabry–Perot resonator;
R_B: reflectivity of back surface of the Fabry–Perot resonator;
$R_\alpha = (R_\mathrm{F}R_\mathrm{B})^{1/2}\mathrm{e}^{-\alpha d}$; $\delta = n(\omega)k_\mathrm{vac}d = n(\omega)\omega d/c$.

One often has $R_\mathrm{F} = R_\mathrm{B} = R$. For this condition and $\alpha d = 0$ one has the simplification A = 1, B = 1.

$$\hat{T} = \frac{1}{1 + F' \sin^2 \delta} , \tag{5.32a}$$

$$\hat{R} = \frac{F' \sin^2 \delta}{1 + F' \sin^2 \delta} = 1 - \hat{T} . \tag{5.32b}$$

with the phase shift δ given in (5.31b) and the *finesse coefficient* F' given by

$$F' = \frac{4R}{(1-R)^2} . \tag{5.32c}$$

A directly related quantity to (a large enough) F' for characterizing a FP resonator is the *finesse* $\mathcal{F}$. This quantity is deduced from considering the transmitted intensity (rather than $\hat{T}$) as a function of frequency ν which is given by the so-called Airy functions. The finesse is then given by the ratio of the *free spectral range* ν_FSR, i.e. the (equidistant) spectral separation of the transmission maxima on the frequency axis, and the linewidth of these maxima $\delta\nu$. The finesse is a measure of the number of rays interfering inside of the FP resonator.

$$\mathcal{F} = \frac{\nu_\mathrm{FSR}}{\delta\nu} = \frac{\pi\sqrt{R}}{1-R} = \frac{\pi}{2}\sqrt{F'} . \tag{5.32d}$$

In Fig. 5.9 we show $\hat{T}$ and $\hat{R}$ as a function of δ for various values of R and αd. For vanishing damping $\hat{T}(\omega)$ reaches unity and $\hat{R}(\omega)$ zero for the conditions of (5.32). Increasing F' makes the FP resonances narrower. $\hat{T}$ and $\hat{R}$ always add up to unity. For finite damping or α this is no longer the case. $\hat{T}$ remains below unity and the height of the resonance decreases with decreasing F', i.e., increasing α for constant values of R. Sometimes lossless, high finesse Fabry–Perot etalons are used as high resolution monochromators. In this case not only the width of the resonance is of importance, but also the free spectral range given in terms of the phase shift by $\delta = \pi$.

A more detailed treatment of the Fabry–Perot and of related problems like dielectric single and multilayer coatings is beyond the scope of this book and [76P1, 85C1, 89T1, 99B1, 16H1] are suggested for further reading.

5.4 Optical Anisotropy

Optical anisotropy arises in solids from the anisotropic nature of the crystal lattice. Consequences are the dependence of the light propagation and absorption on its polarization state or the rotation of the polarization plane during propagation.

5.4.1 Birefringence and Dichroism

Until now we have assumed, for the sake of simplicity, that the solid-state sample is isotropic, i.e., that $\varepsilon(\omega)$ is a scalar function. This has the consequences that considering (2.23), (2.24) and (2.30) $\boldsymbol{D}$ is parallel to $\boldsymbol{E}$ and $\boldsymbol{S}$ is parallel to $\boldsymbol{\Pi}$. Later on we will in general use this assumption again. But now we want to consider the case of optically anisotropic materials. Indeed, many crystals are anisotropic, including the hexagonal wurtzite structure of several technically relevant semiconductors like GaN or ZnO. Even cubic crystals show a weak anisotropy for a finite wave vector $\boldsymbol{k} \neq 0$ since cubic *symmetry* is lower in symmetry than spherical symmetry, see also [16G1]. This latter aspect will not be considered for the moment.

Considering optical excitations as driven oscillators (this model will be discussed in detail in Chap. 7) we can already understand such anisotropies if we assume that some oscillators can be excited (i.e., elongated) only in one direction, e.g., in the x-direction but not in the others. Such an oscillator would react only on the component of an incident electric field polarized $\boldsymbol{E} \parallel \hat{\boldsymbol{x}}$. In the microscopic model the same approach means that the oscillator strength f (introduced also in Chap. 7) depends on the polarization direction of the incident light. It could be, e.g., finite for light polarized parallel to a crystallographic axis and zero perpendicular to it. Indeed it is already sufficient that the oscillator strength is different for different orientations of the polarization with respect to the crystallographic axis in order to obtain the phenomenon of *birefringence*.

To describe the optical anisotropy of a material we recall that the dielectric function $\varepsilon(\omega)$ is generally a tensor. It describes the relation between the two vectors $\boldsymbol{D}$ and $\boldsymbol{E}$:

$$D_i = \varepsilon_0 \sum_j \varepsilon_{ij} E_j; \quad i, j = x, y, z \tag{5.33}$$

In principle the ε-tensor has nine components. It can be shown [74B1, 99B1, 16H1] that the conservation law for the electromagnetic field energy requires $\varepsilon_{ij} = \varepsilon_{ji}$ so we are left with a maximum of six independent components. Furthermore, it can be shown that every symmetric 3×3 tensor can be brought into diagonal form by a suitable rotation of the cartesian coordinate system. If we choose this coordinate system, the $\varepsilon_{ij} = 0$ for $i \neq j$ and we are left with the three elements on the main diagonal $\varepsilon_{xx}, \varepsilon_{yy}, \varepsilon_{zz}$. These diagonal elements are all different in the general case, i.e., for so-called *biaxial crystals*.

Usually one tries to align the cartesian coordinates for $\varepsilon(\omega)$ in a simple way with respect to the crystallographic axes. In *uniaxial crystals* one identifies the z-axis with the crystallographic $\boldsymbol{c}$-axis and the x-y plane with the (usually almost isotropic) plane perpendicular to $\boldsymbol{c}$.

If transformed on these main axes, the tensor $\varepsilon(\omega)$ for a uniaxial system has in the main diagonal two equal elements

$$\varepsilon_{xx}(\omega) = \varepsilon_{yy}(\omega) \neq \varepsilon_{zz}(\omega) \tag{5.34a}$$

and zeros otherwise. For biaxial systems of even lower symmetry, such simple relationships are often no longer possible and one finds

$$\varepsilon_{xx}(\omega) \neq \varepsilon_{yy}(\omega) \neq \varepsilon_{zz}(\omega) \neq \varepsilon_{xx}(\omega) \ . \tag{5.34b}$$

For cubic crystals one finds for $\boldsymbol{k} = 0$: $\varepsilon_{xx}(\omega) = \varepsilon_{yy}(\omega) = \varepsilon_{zz}(\omega)$. This situation allows us to treat $\varepsilon(\omega)$ as a scalar quantity. For $\boldsymbol{k} \neq 0$ birefringence and dichroism (though weak) may appear for certain orientations as already mentioned [81G1, 96T1, 02T1].

Since all of the important semiconductors crystallize either in cubic systems (diamond structure with point-group O_h, zincblende structure T_d or hexagonal ones (wurtzite structure $C_{6\mathrm{v}}$)) we will not go below uniaxial symmetry and refer the reader for these problems to books on crystal optics [76P1, 84A1, 99B1] or on crystallography [12B1]. The properties of "point groups" and the implications of group-theoretical aspects on semiconductor optics will be explained in Appendix B.

Direct consequences of the tensor character of $\varepsilon(\omega)$ are *birefringence* and *dichroism*. Here we want to briefly outline both effects.

Dichroism means literally that a crystal has two different colors depending on the direction of observation. In a more general sense one describes by dichroism every dependence of the absorption spectra on the direction of light polarization. In Fig. 5.10 we show schematically transmission spectra for a dichroic, uniaxial material of a certain thickness. The sample is transparent for both polarizations below $\hbar\omega_1$. The resonance at $\hbar\omega_1$ is assumed to couple more strongly to the light field (i.e., to have larger "oscillator strength") for the polarization $\boldsymbol{E} \perp c$ than for $\boldsymbol{E} \parallel \boldsymbol{c}$. Above $\hbar\omega_2$ light is absorbed almost completely for both polarization orientations. The dichroic region obviously lies between $\hbar\omega_1$ and $\hbar\omega_2$.

In some materials this region covers a wide spectral range, in some cases the whole visible part of the spectrum. In such a case this material can be used as a polarizer. As an example polaroid films contain long organic molecules which have been oriented parallel to each other by the stretching of the film during the manufacturing process. These molecules absorb radiation polarized parallel to the chain, and transmit for the perpendicular orientation over most of the visible spectrum. Another material which is known for its dichroism is tourmaline which comes in differently colored varieties. In many semiconductors the dichroic spectral range is rather narrow and amounts often only to a few tens of meV. These materials are, of course, of no use as commercial polarizers, but the investigation of the dichroism gives very important

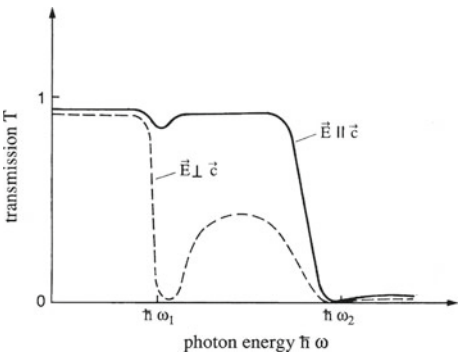


Fig. 5.10 Schematic drawing of the transmission spectra of a dichroic material for two different polarizations of light with respect to the crystallographic axis

information on the symmetries and selection rules of the resonances. We will see some examples of this effect in Chap. 21.

If we assume that the eigenfrequencies and/or oscillator strengths of some resonance(s) depend on polarization, then we know immediately from the Kramers–Kronig relations (see Appendix. A), that the real part of the refractive index $n(\omega)$ also depends on the orientation of $\boldsymbol{E}$ relative to $\boldsymbol{c}$. This means that the material is birefringent. We can even state that every dichroic material must show birefringence and that birefringent materials must have some spectral range in which dichroism occurs.

For uniaxial materials (e.g., crystals with uniaxial symmetry C_{6v} or as a prototype calcite with point group D_{3d}) an incident light beam can always be decomposed into two components of the electric field polarized parallel and perpendicular to the *main section*. The main section is the plane defined by the crystallographic axis and the incident wave vector. The beam polarized perpendicular to the main section is called the *ordinary* (o) *beam*. Its refraction is described by Snell's law and the refractive index $\tilde{n}(\omega)$ is independent of polarization orientation. This fact can be understood since the ordinary beam is always polarized perpendicular to the $\boldsymbol{c}$-axis and we assume that uniaxial materials are isotropic in the plane $\perp \boldsymbol{c}$. This is strictly correct only for vanishing wave vectors (and corresponds just to the situation for the dipole approximation in Sect. 6.2) and to a very good approximation for small but finite $\boldsymbol{k}$ values.

The so-called *extraordinary* (eo) *beam*, the polarization of which falls in the main section, has components $\boldsymbol{E} \parallel \boldsymbol{c}$ and $\boldsymbol{E} \perp \boldsymbol{c}$ the weights of which depend on the angle $\gamma = \angle(\boldsymbol{k}, \boldsymbol{c})$. It is not surprising that the refractive index experienced by the extraordinary beam depends on γ, since the relative coupling to the oscillators active for the orientations $\parallel \boldsymbol{c}$ or $\perp \boldsymbol{c}$ changes with the angle γ.

For a general direction of incidence an unpolarized (or elliptically polarized) beam will be decomposed into two beams polarized perpendicular to each other—the ordinary and the extraordinary ones—which will be separated in space, as shown schematically in Fig. 5.11. This is the concept, which allows us to use birefringent materials as polarizers. We note here already that the wave vector $\boldsymbol{k}$ and the Poynt-

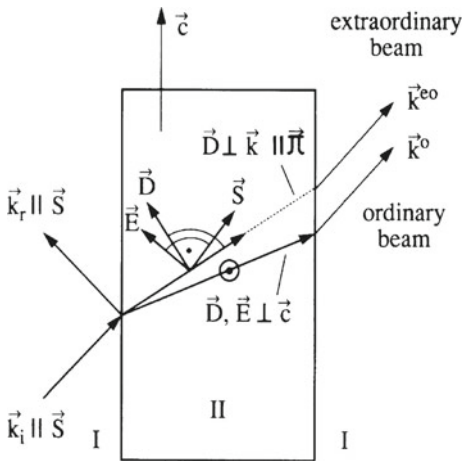


Fig. 5.11 Birefringence for an unpolarized beam falling on a birefringent material at a given angle with the crystallographic axis $\boldsymbol{c}$ parallel to the interface

ing vector $\boldsymbol{S}$ of the eo beam are not necessarily parallel to each other. The reason will be given in connection with Fig. 5.13. In the case of Fig. 5.11 the refractive index of the ordinary beam is greater than that of the extraordinary one. This situation is called *negative-uniaxial birefringence*, the opposite situation correspondingly *positive-uniaxial birefringence*, since Δn is defined as

$$\Delta n := n_{\mathrm{eo}} - n_{\mathrm{o}} \tag{5.35}$$

where the indices o and eo stand for ordinary and extraordinary.

There are two limiting orientations which result in rather clear and simple situations. Therefore these orientations are usually investigated in semiconductor optics. One situation is $\boldsymbol{k} \parallel \boldsymbol{c}$. In this case the $\boldsymbol{E}$ field can be only perpendicular to $\boldsymbol{c}$, this means one observes the ordinary beam only, independent of the polarization of the incident beam. The other clear orientation is $\boldsymbol{k} \perp \boldsymbol{c}$. In this case one can choose by a polarizer the orientation $\boldsymbol{E} \perp \boldsymbol{c}$ for the ordinary beam or $\boldsymbol{E} \parallel \boldsymbol{c}$ for the extraordinary beam. In the latter situation, the $\boldsymbol{E}$ field acts only on oscillators which can be elongated parallel to $\boldsymbol{c}$.

Oblique incidence on a surface cut parallel or perpendicular to $\boldsymbol{c}$ or normal incidence on a surface cut under an arbitrary angle (Fig. 5.12) with respect to $\boldsymbol{c}$ are much more complicated to evaluate concerning the spectra of reflection or transmission. The worst situation is, of course, oblique incidence on a plane at an arbitrary angle to the $\boldsymbol{c}$-axis. Scientists working on semiconductor optics usually try to avoid these situations, scientists devoting their work to crystal optics find it challenging and even prefer biaxial systems (see below with Fig. 5.15) to others.

A scenario which allows us to discuss various aspects of birefringence is perpendicular incidence on a plane at an oblique angle with respect to $\boldsymbol{c}$. This situation is shown in Fig. 5.12. The incident beam is split into two when entering the birefringent material. The ordinary beam continues to propagate normal to the surface as

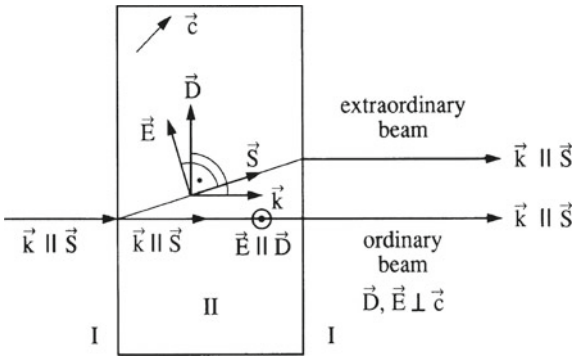


Fig. 5.12 Same as Fig. 5.11 but for normal incidence on a surface cut under an arbitrary angle with respect to $\boldsymbol{c}$

expected from Snell's law (5.11a) or from the conservation of momentum parallel to the surface (5.15). The extraordinary one seems to violate these two rules. Since a violation of the law of momentum conservation would be very serious, not only for physics, we have to look more closely at this problem.

There are two rather simple ways to present birefringence. One, which we shall outline towards the end of this subsection is in terms of the *indicatrix*. In the other method, one uses in polar coordinates a plot which gives the phase velocity $v_{\mathrm{ph}} = cn^{-1}(\omega)$. This is basically the inverse of the real part of $\tilde{n}(\omega)$ as a function of the direction of propagation. In a uniaxial system, the figures produced when we include all directions are a sphere for the ordinary beam and a figure with rotational symmetry with respect to the $\boldsymbol{c}$-axis for the extraordinary one. They touch for propagation of light parallel to the crystallographic axis as shown in Fig. 5.13 since there are only o-waves for this orientation, as mentioned above. Now we recall Huygens' principle for the propagation of light which says that every point illuminated by an incident primary wave front becomes the source of a secondary wavelet such that the primary wavefront at a later time results from the superposition of the amplitudes of all wavelets. In addition, we must discard the back-travelling waves in the way shown by Fresnel and Kirchhoff which would appear otherwise from the above given principle. With this amendment we can construct wavefronts when we identify the shape of the wavelets with Fig. 5.13. For more details see e.g. [96T1, 02T1, 12B1] and references given therein.

Shown in Fig. 5.12 are the $\boldsymbol{S}$, $\boldsymbol{k}$, $\boldsymbol{E}$ and $\boldsymbol{D}$ vectors. The vectors $\boldsymbol{k}$ and $\boldsymbol{\Pi}$ are always parallel to each other. The magnetic vectors $\boldsymbol{B}$ and $\boldsymbol{H}$ are parallel to each other and normal to the electric vectors and are, in this context, of no further interest since we are dealing with nonmagnetic material.

In Fig. 5.14a the same situation is shown for the o-beam. The wavelets are spheres, the resulting wave front is parallel to the vacuum-medium interface. The vectors of energy-flux density $\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H}$, of momentum density $\boldsymbol{\Pi} = \boldsymbol{D} \times \boldsymbol{B}$ and the wavevector $\boldsymbol{k}$ ($\hat{=}$ momentum $\hbar\boldsymbol{k}$) are parallel. (Diffraction effects caused by the finite beam diameter are neglected here, though they are obviously also described by Huygens' principle.) The situation for the eo-beam is presented in Fig. 5.14b. The

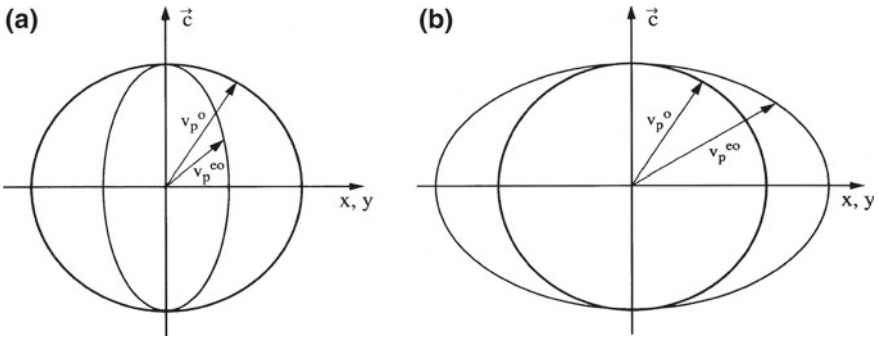


Fig. 5.13 The phase velocity of light in an optically uniaxial material shown for one frequency ω in polar coordinates for positive (**a**) and negative (**b**) birefringence for the ordinary (*o*) and extraordinary (*eo*) beams

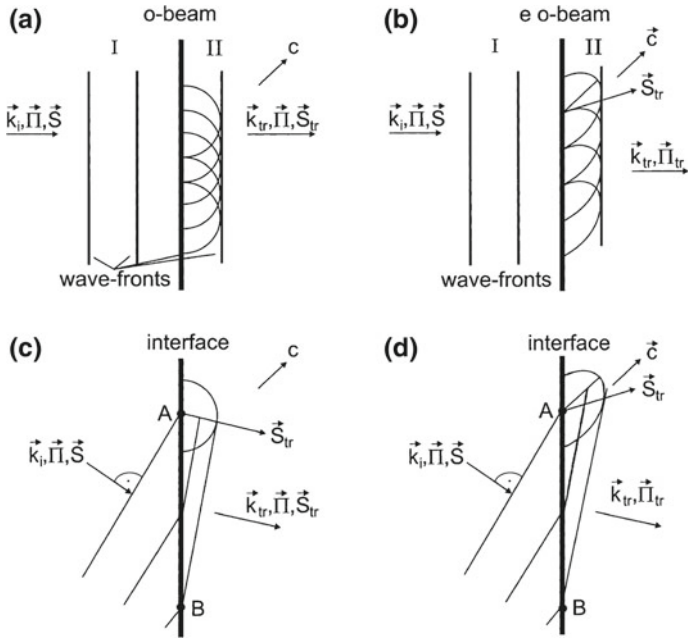


Fig. 5.14 The construction of the wave fronts of the ordinary (**a**, **c**) and the extraordinary (**b**, **d**) beams for various geometries using Fig. 5.13

wave front produced by the superposition (or interference) of the wavelets, and constructed as the tangent to the wavelets, is still parallel to the interface. This wave front describes the $\boldsymbol{D}$ field because we know from Maxwell's equations and the boundary conditions of Sect. 5.1.1 that the boundary condition for $\boldsymbol{D}$ is that the normal component D_n is continuous over the interface. Here, $D_\mathrm{n} = 0$ on both sides. As a

consequence, the classical momentum density $\boldsymbol{\Pi}$ and the momentum $\hbar\boldsymbol{k}$ of the light quanta are still perpendicular to the interface, as required by the conservation of the momentum component parallel to the interface, which is here obviously zero.

On the other hand, we can see that the whole wave front is shifting sideways with continuing propagation into the medium. This shift is described by the Poynting vector $\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H}$. The direction of this vector is given by the origin of the wavelet and the point where the tangent touches it. The $\boldsymbol{E}$ field is necessarily perpendicular to $\boldsymbol{S}$. As required by the boundary condition for $\boldsymbol{E}$ as deduced from Maxwell's equations, the tangential component $\boldsymbol{E}_\mathrm{t}$, must be the same on both sides of the interface (including incident, refracted and reflected beams). The normal component of $\boldsymbol{E}$ can change, and that is what happens in the orientation of Fig. 5.14b.

To summarize, we observe that there are no violations of any conservation laws. The tangential components of the momentum are conserved at the interface and for this quantity Snell's law is still valid. However, the direction of energy propagation given by $\boldsymbol{S}$ changes, but there are no conservation laws for this direction. The law of energy conservation itself has, of course, to be fulfilled, this means in this case that the total amount of energy falling per unit time on the interface equals the sum of transmitted and reflected energies.

Figure 5.14c, d finally show schematically the general situation i.e., oblique incidence on a surface cut at an arbitrary angle with respect to $\boldsymbol{c}$. The construction simply uses the size of the wavelets around point A when the incident wave front just reaches the interface at point B. We see that the refractive index for $\boldsymbol{k}$ is different for both polarisations and that the directions of $\boldsymbol{k}_\mathrm{tr}$ and $\boldsymbol{S}_\mathrm{tr}$ differ for the eo beam.

The basic idea of the representation of birefringence by the indicatrix is explained in connection with Fig. 5.15. We plot from the origin lines in all directions with a length equal to the refractive index $n(\omega)$ of a wave with $\boldsymbol{D}$ polarized in this direction. In doing so, we get a sphere for an isotropic material (Fig. 5.15a) and ellipsoids with rotational symmetry for uniaxial materials. In Figs. 5.15b,c we show the situation for positively and negatively birefringent uniaxial crystals, respectively. The rotation axis coincides in this case with the crystallographic $\boldsymbol{c}$-axis. For crystals of even lower symmetry, all three axes of the indicatrix have different lenghts (Fig. 5.15d) in agreement with the situation for $\varepsilon(\omega)$ in (5.34b).

The use of the *indicatrix* is now the following: For a given direction of propagation $\boldsymbol{k}$ one cuts the indicatrix with a plane normal to $\boldsymbol{k}$ which contains the origin. In general, this cross section is an ellipse except for isotropic materials. The distance from the origin to this ellipse in a certain direction gives the refractive index of a beam with $\boldsymbol{D}$ polarized in this direction and propagating in the direction of $\boldsymbol{k}$. The optical axes are for anisotropic materials (Fig. 5.15b–d) now defined as those directions of propagation for which the cross section of the indicatrix is a circle. For the uniaxial materials with an indicatrix of rotational symmetry around $\boldsymbol{c}$ in Fig. 5.15b,c there is obviously only one optical axis, which coincides with the $\boldsymbol{c}$-axis as stated already above. For systems of lower symmetry, where all three main axes of the indicatrix (and all three $\varepsilon_{ii}(\omega), i = x, y, z$ of the dielectric tensor) are different, one finds two optical axes, which generally do not coincide with any of the cartesian coordinates (Fig. 5.15d) and the direction of which can even vary with ω.

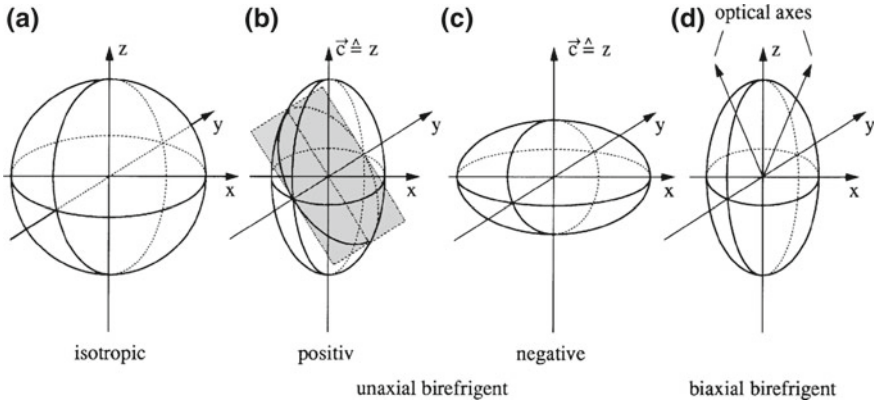


Fig. 5.15 The indicatrix for isotropic (**a**) uniaxial (**b**, **c**) and biaxial materials (**d**)

It should be mentioned that birefringence can be induced in non-birefringent materials, e.g., by the application of electric or strain-fields. These phenomena are known as *Pockels* and *Kerr effect* and as mechanical or stress birefringence or *photoelasticity*, respectively. The Pockels effect occurs in crystals without centers of inversion and the effects vary proportionally with $\boldsymbol{E}$. The Kerr effect occurs in isotropic solids like glass, in fluids like nitrobenzene and varies therefore proportionally to $\boldsymbol{E}^2$. While these effects are often deliberately investigated or exploited (e.g. in modulators), especially the stress induced birefringence can produce unwanted, spurious effects in semiconductor optics if a sample is not mounted strain-free, e.g., in a cryostat.

All the phenomena of birefringence which we discussed here for a uniaxial material are as mentioned at the beginning of this subsection, simply due to the fact that the dielectric function $\varepsilon(\omega)$ is a tensor for crystallographic systems of symmetry lower than cubic and that $\boldsymbol{D}$ and $\boldsymbol{E}$ need therefore no longer be parallel to each other. More information on the topic of this subsection is found in [76P1, 84A1, 96T1, 99B1, 02T1, 07M1, 07S1, 09R1, 12B1, 16H1] and references therein.

5.4.2 Optical Activity

The last aspect we want to mention in this chapter on macroscopic optical properties is *optical activity*. In this case of optical anisotropy a linearly polarized light beam remains linearly polarized when propagating through an optically active medium but the direction of polarization rotates slowly by an amount δ_{rot} that varies with increasing distance l, which is the distance the light travels through the medium, i.e. the rotatory power $\varrho(\omega)$

$$\delta_{\text{rot}} = \varrho(\omega) l \tag{5.36a}$$

where $\rho(\omega)$ depends on the material and on ω.

While a birefringent material splits an incident beam in two, orthogonally and linearly polarized ones, which propagate with different phase velocities through the sample and thus accumulate a phase shift with increasing distance (see Problem 5.7), an optically active material splits a linearly polarized incident beam into a (σ^-) (spin vector pointing along $-\boldsymbol{k}$) and (σ^+) (spin vector pointing along $+\boldsymbol{k}$) circularly polarized one with equal amplitude. Again both components travel with different phase velocities. A superposition after a certain distance results again in linearly polarized light as long as there is no difference in the absorption for σ^+ and σ^- polarized light, but with a tilted direction of polarization.

The rotatory power $\varrho(\omega)$ is with (5.36a) equal to the rotation angle per unit length of material and is given by

$$\varrho(\omega) = \frac{\pi}{\lambda_0}(n_- - n_+) \tag{5.36b}$$

It can be described by an $\varepsilon(\omega)$ tensor of the form

$$\begin{pmatrix} D_1 \\ D_2 \\ D_3 \end{pmatrix} = \varepsilon_0 \begin{pmatrix} n^2 & -iG & 0 \\ iG & n^2 & 0 \\ 0 & 0 & n^2 \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \\ E_3 \end{pmatrix} \tag{5.36c}$$

where the real parts of the refractive indices $n_\pm$ for the two circular polarisations are given by

$$n_\pm = \sqrt{n^2 \pm G} \tag{5.36d}$$

This phenomenon occurs in amorphous solids, in liquids and in gases if they contain molecules with a well defined chirality or handiness like dextrose or laevulose which have both the chemical composition $C_6H_{12}O_6$. In crystalline solids, including semiconductors, optical activity occurs if either the molecules in the unit cell have a chirality or if atoms or isotropic molecules are arranged in the unit cell in such a way that a screw-axis appears. Examples are quartz or to a minor extend wurtzite type crystals [79I1] parallel to their c-axes, which are both essentially built from tetrahedrons, but are arranged in a way that a shift and a rotation along the axis transforms the crystal into itself.

The optical activity induced in a material by a magnetic field in the direction of light propagation is known as the *Faraday effect* described by the *Verdet constant* (better Verdet material parameter) $V(\omega)$ via

$$S_+ = V(\omega)\, l\, B\ . \tag{5.36e}$$

The phenomenon of *magnetic dichroism* occurs when a difference in the absorption spectra for σ^+ and σ^- polarized light is induced by the B-field.

Problems

5.1 Consider the interface between vacuum (or air) and glass ($n = 1.45$) at a wavelength λ_{vac} of 0.5 μm. Calculate for an angle of incidence $\alpha_i = 45°$ the incident, reflected, and transmitted wave vectors, and the transmitted and reflected intensities for both polarizations. Calculate Brewster's angle for the transition air→glass and glass→air and the angle for the onset of total internal reflection.

5.2 Find a piece of polarizing material (polaroid) and observe the light reflected from a nicely polished floor or scattered from the blue sky using different orientations of the light propagation and of the polarization. Do not look into the sun! Try to explain your findings.

5.3 From Figs. 5.11 and 5.12 one can understand that a birefringent crystal can be used as a polarizer if the lateral diameter of the beams is smaller than their lateral displacement after the passage. Usually one uses slightly more complex arrangements known as Glan–Thomson or Taylor polarizers. Make yourself familiar with the way of operation of these optical components. What can happen if you use them under oblique incidence or with a di- or convergent light beam with large angle of aperture?

5.4 Derive the laws of reflection and refraction from the principle of Fermat (see Fig. 5.4) and from momentum conservation (5.16). Does this law also hold if you kick a soccer ball against a wall?

5.5 Consider Fig. 5.9c. Explain in words why only in this diagram the three curves intersect, i.e., why zero absorption gives both the highest maxima and the lowest minima in reflectance.

5.6 Verify the energy conservation law (5.23a) with the help of (5.23b) and (5.18).

5.7 What is a quarter ($\lambda/4$) or half ($\lambda/2$) wave plate? How thick is it? What is the state of polarisation of the transmitted light beam, when the incident beam is linearly polarized? Does it depend on the orientation of the incident linear polarization with respect to the crystallographic axis. Make a simple sketch. (Generally one uses a uniaxial material and normal incidence for theses devices.) What is a low order $\lambda/4$ plate? Why are achromatic $\lambda/4$ plates rather expensive?

5.8 Make a simple sketch to make yourself familiar with optical activity. Can you imagine, that the absorption can be different for σ^+ and σ^- polarized light?

5.9 Verify some of the prominent features in Fig. 5.6 with (5.18), e.g., normal and grazing incidence, α_B or α_{tot}.

References

[74B1] G.L. Bir, G.E. Pikus, *Symmetry and Strain Induced Effects in Semiconductors* (Wiley, New York, 1974)
[75H1] H.C. van de Hulst, *Light Scattering by Small Particles* (Wiley, New York, 1957)
[76P1] R.W. Pohl, *Optik und Atomphysik*, 13th edn. (Springer, Berlin, 1976)
[79I1] El. Ivchenko, A.V. Sel'kin, JETP **76**, 1837 (1979)
[81G1] O.V. Gogolin et al., JETP Lett. **34**, 328 (1981)
[84A1] V.M. Agranovich, V.L. Ginzburg, *Crystal Optics with Spatial Dispersion and Excitons*, vol. 42, 2nd edn., Springer Series Solid-State Science (Springer, Berlin, 1984)
[85C1] D. Craig et al., IEEE J. **QE-21**, 1363 (1985)
[89T1] A. Thelen, *Design of Optical Interference Coatings* (McGraw Hill, New York, 1989)
[96T1] E. Tsitsishvili, Appl. Phys. A **62**, 255 (1996); Solid State Commun. **100**, 541 (1996)
[99B1] M. Born, E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, 7th edn. (Cambridge University Press, Cambridge, 1999)
[00W1] W.H. Weber, R. Merlin (eds.), *Raman Scattering in Materials Science*, vol. 42, Springer Series in Materials Science (Springer, Berlin, 2000)
[02T1] E. Tisitishvili, *Spectroscopy of Systems with Spatially Confined Structures* (2001). NATO Science Series II, vol. 90 (Kluwer, Dordrecht, 2002), p. 357
[07M1] D. Meschede, *Optics, Light, and Lasers*, 2nd edn. (Wiley, Weinheim, 2007)
[07S1] B.E.A. Saleh, M.C. Teich, *Fundamentals of Photonics*, 2nd edn. (Wiley, Hoboken, 2007)
[09R1] H. Römer, *Theoretical Optics*, 2nd edn. (Wiley-VCH, Weinheim, 2009)
[12B1] W. Borchardt-Ott, *Chrystallography*, 3rd edn. (Springer, Berlin, 2012)
[16G1] M. Grundmann, *The Physics of Semiconductors*, 3rd edn. (Springer, Berlin, 2016)
[16H1] E. Hecht, *Optics*, 5th edn. (Pearson, New York, 2016)

Chapter 6
Microscopic Aspects of Weak Light-Matter Coupling

Abstract The preceding chapter has described the macroscopic properties of light-matter interaction. In this chapter we present the basic interaction processes between light and matter from a microscopic point of view. We use here the perturbative or weak-coupling approach for this interaction, which is in most cases sufficient for dilute systems such as gases. For solids the strong-coupling approach is often necessary. This leads to the concept of polaritons, which is introduced later in Chap. 8. We will first outline the basic interaction mechanisms between light and matter, namely absorption, spontaneous and stimulated emission. Then we proceed to the treatment of linear optical properties in the framework of perturbation theory and derive Fermi's golden rule.

6.1 Absorption, Stimulated and Spontaneous Emission, Virtual Excitation

For simplicity we assume that we have a certain number of *two-level system* as shown in Fig. 6.1. Each of these "atoms" has one electron which can be either in the ground or in the excited state. Later we will extend the model from such a two-level system to bands in semiconductors, but the basic interaction processes remain the same.

In Fig. 6.1a an incident photon hits an atom in its ground state. With a certain probability the photon is annihilated and the electron gains enough energy to reach the excited state. For reasons of energy conservation, the photon has to fulfill the condition

$$\hbar\omega = E_{\mathrm{ex}} - E_{\mathrm{g}} \ , \tag{6.1}$$

where $E_{\mathrm{ex}} - E_{\mathrm{g}}$ is the energy difference between the ground and excited states. We call this process *absorption* in agreement with the definition in Sect. 5.2 if the energy of the photon is soon converted into other forms of energy, that is, if the electron undergoes some scattering processes, which destroy its coherence or more precisely the coherence of the electric polarization connected with this transition to the incident light field (the aspects of coherence will be treated in more detail in Part B of this book. The electron eventually returns to its ground state and looses its energy e.g.,

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_6

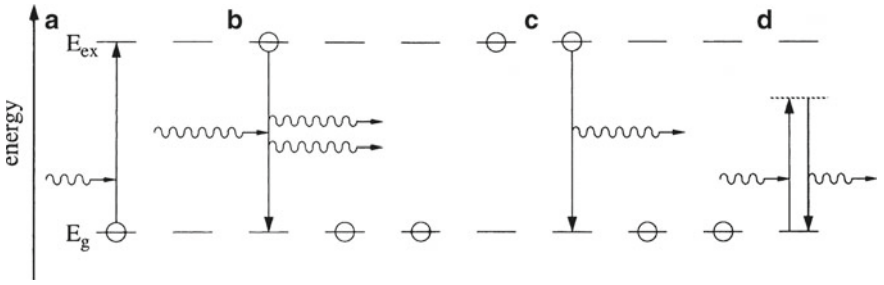


Fig. 6.1 Schematic representations of various interaction processes of light with matter namely absorption (**a**), stimulated emission (**b**), spontaneous emission (**c**) and virtual excitation (**d**)

to the phonon system, i.e., as heat or as a photon which is not coherent with the incident one. The first process is called *non-radiative recombination*, the second one *resonance fluorescence*.

If an incident photon hits an atom with its electron in the excited state it can induce with a certain probability a transition of the electron from the excited to the ground state (Fig. 6.1b). In this process a second photon is created which is identical in momentum, energy, polarization and phase to the incident one. This process is called *induced* or *stimulated emission*. This process can be used to amplify a photon field. It is therefore the basic mechanism for all lasers (Light Amplification by Stimulated Emission of Radiation). Absorption and stimulated emission are both induced and thus closely related events.

An electron in the excited state can also with a certain probability return to the ground state by itself, either by emitting a photon (Fig. 6.1c) or by loosing the transition energy through phonons or collisions. In the present context, the first mechanism is of interest. It is called *spontaneous emission* or spontaneous *radiative recombination*, while the second possibility is known as *non-radiative recombination*.

Spontaneous emission can also be understood in a different way. In Chap. 3 we saw in connection with (3.7) that photons are similar to harmonic oscillators and thus have a zero-point energy. This zero-point energy exists for all photon modes. It cannot be absorbed because a harmonic oscillator does not have any states below the zero-point energy, but it can induce a transition in the way as discussed in connection with Fig. 6.1b. So we can consider spontaneous emission as a process induced by the zero-point vibrations of the electro-magnetic field, which are also called vacuum fluctuations or zero point fluctuations of the electro-magnetic field.

The last process presented here is *virtual excitation*. Understanding this phenomenon often causes some problems for students. So we develop this topic slowly and try to explain it from various points of view in context with Fig. 6.1d. Virtual excitation means the creation of a state with the same wave function as the excited state, but with an energy which is different from the eigenenergy of this excited state. This process becomes possible through the *uncertainty principle* of quantum mechanics which can be written in space and momentum coordinates:

$$\Delta x_i \, \Delta p_i \geq h \quad \text{for} \quad i = 1, 2, 3 \; . \tag{6.2a}$$

A similar relation exists for energy and time

$$\Delta E \, \Delta t \approx \hbar \; . \tag{6.2b}$$

The latter relation means that it is possible to violate energy conservation by an amount ΔE up to a maximum time Δt which fulfills the above condition. Or, in other words, if we want to define the energy of a state with a certain precision ΔE, this state has to exist at least for a duration of Δt. In principle (6.2b) is valid also in simple classical wave theory (e.g., acoustics) and is very well known from Fourier transformations. A harmonic oscillation with central frequency ω which lasts only for a time Δt, has a spectral width $\Delta\omega$ given by

$$\Delta\omega \Delta t \geq 1 \; . \tag{6.2c}$$

The connection between (6.2b) and (6.2c) comes simply from the relation

$$E = \hbar\omega \; . \tag{6.2d}$$

It should be mentioned that a similar set of arguments holds for (6.2a), too.

If we send now a photon with energy $\hbar\omega'$ to the atom, we can excite the electron for a maximum time Δt given by (6.2b) or for our specific case by

$$\Delta t \approx \hbar \, | \, (E_{\text{ex}} - E_{\text{g}}) - \hbar\omega' \, |^{-1} \quad . \tag{6.3}$$

At the latest after the time Δt has elapsed, the excited state must collapse. The simplest way to do this is to emit a photon identical to the one which caused the virtual excitation. This "new" photon has however a certain phase delay with respect to the incident photon, because the energy was stored for a maximum time Δt in the atom. As a consequence, an electromagnetic wave propagates with a lower phase-velocity through an ensemble of atoms than through vacuum. The same effect is described in a phenomenological way by the refractive index $n(\omega)$ (compare (2.18) and (2.44). So we get a first hint of how $n(\omega)$ can be understood and calculated in quantum mechanics. Obviously Δt increases if we approach the resonance condition in (6.3) and consequently wave propagation through a material will deviate more strongly from that through vacuum. This is indeed the case, as we shall see in Sect. 7.3.

If the virtually excited state emits a photon $\hbar\omega'$ in a direction different from the incident one, we have a scattering process as discussed in connection with (5.26). If $\hbar\omega'$ approaches the resonance energy, this scattering process is also known as resonance fluorescence as mentioned already above.

In connection with this scattering, we may ask how can light propagate at all in a clear or transparent medium? The answer is that we have in dense media many scattering atoms or centers in the coherence volume of light, independent of whether it is light from an incandescent lamp, a laser or another source. As a consequence,

every scattered wave finds another one which has a phase difference, resulting in destructive interference. The only way that all scattered waves interfere constructively is just the usual propagating wave. The explanation for the phenomenon of a blue sky, which we gave in Sect. 5.2, is consistent with this picture and results from the following conditions. Here the diameter of the scattering centers is small compared to the wavelength of the scattered radiation. But now, there are only a few scattering centers, i.e., gas molecules, density fluctuations or other inhomogenities per coherence volume of sunlight, so that the mutually destructive interference of the scattered waves is not complete. For a detailed discussion of spatial and temporal *coherence* see e.g., [93S1, 07M1, 07S1] and second volume of this book.

If the virtually excited state disappears under simultaneous emission of a photon and the creation or annihilation of a phonon (i.e., a quantum of the lattice vibrations, as in Chap. 11), energy conservation for the emitted photon $\hbar\omega_R$ implies

$$\hbar\omega_R = \hbar\omega' \pm \hbar\Omega_{\text{phonon}} \ . \tag{6.4}$$

This phenomenon is called *Raman scattering* for optical phonons, and *Brillouin scattering* for acoustic phonons. The "−" sign gives the Stokes and the "+" sign the anti-Stokes emission. Similar processes are also possible with more than one phonon or with excitations in the electronic or spin system of the semiconductor.

Already from the few phenomena outlined briefly above which involve virtual excitation, it is obvious that this mechanism is of some importance for the optical properties of matter. Therefore we want to examine it from another point of view and outline the well-known classical analog of the virtually excited states. In addition this analogy gives some justification for the calculation of the dielectric function used in Chap. 7.

Virtual excitation in quantum mechanics corresponds to a driven or forced oscillation in classical mechanics. If we have an oscillator of eigenfrequency ω_0 (corresponding to the energetic differences $E_{ex} - E_g$ in quantum mechanics) and if we excite it with an external frequency ω, it will oscillate with frequency ω after a short damped transient feature of oscillations with ω_0. The amplitude of these steady oscillations increases with decreasing detuning $|\omega - \omega_0|$ depending on the properties of the oscillator, e.g., its damping. This increase of the amplitude of the classical oscillator corresponds to the increase of Δt in (6.3) in the picture of virtual excitation. It can be understood qualitatively in the sense that the strongest deviations of $\varepsilon(\omega)$ or $\tilde{n}(\omega)$ from the vacuum value $\varepsilon = \tilde{n} = 1$ occur in the vicinity of the resonance ω_0. We will elaborate this concept in detail in Chap. 7.

However, before doing so, we shall demonstrate how the various transitions shown in Fig. 6.1 and some others can be treated quantitatively in the limit of weak coupling by *perturbation theory*.

6.2 Perturbative Treatment of the Linear Interaction of Light with Matter

The Hamiltonian of the total system consisting of the electron states in an ensemble of two-level atoms (see Fig. 6.1) (or of the bands of a semiconductor), the radiation field and the interaction of these two quantum systems can be written as

$$H = H_{\mathrm{el}} + H_{\mathrm{rad}} + H_{\mathrm{interact}} \ . \tag{6.5}$$

In the picture of second quantization outlined in Sect. 3.2 H_{interact} contains terms which describe, e.g., the annihilation of a photon and of an electron in the ground state and the creation of an electron in the excited state for the process shown in Fig. 6.1a, weighted with a factor which contains the transition matrix element.

The exact solution of the total Hamiltonian leads to the *polariton* concept (Chap. 8). Here the coupling between the two-level atoms and the light field is coherent. This results in a mixing of these entities accompanied by a level splitting in analogy to the case of two coupled oscillators. We will elaborate this quasi-classical concept of *strong coupling* in Chap. 8. It not only describes many phenomena in linear solid-state optics but also in quantum optics in which we will delve into in the second volume of this book.

Often however, the mentioned coherence is rapidly destroyed e.g., by scattering. The induced life-time broadening of the resonances is then much larger than the resonance splitting. I.e., the coupling is essentially incoherent and the two quantum systems are treated as independent entities. In this *weak-coupling regime* one can treat the radiation field as a small perturbation, that is, we assume that the eigenstates φ_{n} and the eigenenergies E_{n} of H_{el}, essentially do not change in the presence of the electromagnetic field, and that the eigenstates of H_{rad} are the photons described already in Chap. 3.

The approximation which we use now is known as the *semi-classical treatment of radiation*. It consists of replacing the canonical conjugate momentum $\boldsymbol{p}$ in the Hamilton function by

$$\boldsymbol{p} \rightarrow \boldsymbol{p} - e\boldsymbol{A} \ , \tag{6.6}$$

where $\boldsymbol{A}$ is the vector potential (3.1). If we replace $\boldsymbol{p}$ by its operator

$$\boldsymbol{p} = \frac{\hbar}{\mathrm{i}} \nabla \tag{6.7}$$

the single-particle Hamiltonian reads

$$H = \frac{1}{2m} \left(\frac{\hbar}{\mathrm{i}} \nabla - e\boldsymbol{A} \right)^2 + V(\boldsymbol{r}) \tag{6.8}$$

including any electrostatic potential into $V(\boldsymbol{r})$. Making use of the Coulomb gauge (3.2) we can evaluate (6.8) to obtain

$$H = -\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{r}) - \frac{e}{m}\boldsymbol{A}\frac{\hbar}{\mathrm{i}}\nabla + \frac{e^2}{2m}\boldsymbol{A}^2 \ , \tag{6.9}$$

$$H = H_{\mathrm{el}} - \frac{e}{m}\boldsymbol{A}\frac{\hbar}{\mathrm{i}}\nabla + \frac{e^2}{2m}\boldsymbol{A}^2 \ , \tag{6.10}$$

$$H = H_{\mathrm{el}} + H^{(1)} + H^{(2)} \ . \tag{6.11}$$

In (6.11) there are two perturbation terms $H^{(1)}$ and $H^{(2)}$. If we assume that A and thus the light intensity are small, and in the regime of linear optics they are small by definition, then $H^{(1)}$ is a perturbation term of first order and $H^{(2)}$ is small of second order. Consequently $H^{(1)}$ has to be used in first-order perturbation theory. In the second-order approximation we have to use $H^{(1)}$ in second-order and $H^{(2)}$ in first-order perturbation theory, etc. We shall come back to this latter aspect in the second volume of this book.

We now want to derive an important relation called *Fermi's golden rule* from which one can calculate the transition rate W_{fj} from an ensemble of initial states i (e.g., the ground states g in Fig. 6.1) to an ensemble of final states f (e.g., the excited states ex in Fig. 6.1). We start with the time-dependent Schrödinger equation

$$H\psi = \mathrm{i}\hbar\frac{\partial\psi}{\partial t} \tag{6.12a}$$

with the stationary solutions for the unperturbed situation $H = H_0$

$$\psi_n(\boldsymbol{r}, t) = \varphi_n(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}(E_n/\hbar)t} \ . \tag{6.12b}$$

For the solution in the presence of a perturbation $H^{(1)}$ we make the ansatz

$$\psi(\boldsymbol{r}, t) = \sum_n a_n(t)\varphi_n(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}(E_n/\hbar)t} \ . \tag{6.13}$$

We assume that the perturbation is switched on at $t = 0$. Before this time the system is in state i, i.e.,

$$\left.\begin{aligned} a_i(t) &= 1 \\ a_{n\neq i}(t) &= 0 \end{aligned}\right\} \quad \text{for } t \leq 0 \ . \tag{6.14}$$

For $t > 0$ the $a_{n\neq i}(t)$ start to grow and under these conditions the transition rate W_{fi} of *Fermi's golden rule* becomes

$$W_{fi} = \frac{2\pi}{\hbar}\left|H_{fi}^{(1)}\right|^2 D(E) \ , \tag{6.15}$$

Here $D(E)$ is the density of the final states modified by momentum conservation if applicable. $H_{fi}^{(1)}$ is the transition matrix element given by

$$H_{fi}^{(1)} = \int \psi_f^*(\boldsymbol{r}) H^{(1)} \psi_i(\boldsymbol{r})\, d\tau =: < \psi_f \mid H^{(1)} \mid \psi_i > \,. \tag{6.16}$$

For a non-degenerate two level system $D(E)$ is simply one per atom. The square of the transition matrix element $\left|H_{fi}\right|^2$ is known as the *transition probability*. Later on we will assume, for simplicity of writing, that some constant factors as the term $2\pi/\hbar$ are incorporated in this $\left|H_{fi}\right|^2$. Transition probabilities are given apart from some coefficients by the square of the respective transition matrix elements of (6.16) in the case of first order perturbation and by the terms (6.17) for second order.

The transition rate W_{fi} is proportional to the transition probability multiplied by the square of the amplitude of the perturbation $H^{(1)}$, i.e., here by $|\boldsymbol{A}_0|^2 \sim I$ where I is the light intensity, i.e., the energy flux per unit of area and time.

If the first-order perturbation term (6.16) vanishes, then according to what we stated above the second-order contribution reads

$$w_{fi}^{(2)} = \frac{2\pi}{\hbar} \left| \sum_{k \neq i,\, f} \frac{H_{fk}^{(1)} H_{ki}^{(1)}}{E_i - E_k} + H_{fi}^{(2)} \right|^2 D(E) \,. \tag{6.17}$$

We restrict ourselves for the moment to the first order according to (6.16) and discuss the term $H_{fi}^{(1)}$ in some more detail for the perturbations of (6.10). With the vector potential $\boldsymbol{A}$ chosen to be harmonic

$$\boldsymbol{A} = \boldsymbol{A}_0 e^{\mathrm{i}(\boldsymbol{kr} - \omega t)} \tag{6.18}$$

we find, e.g., for the absorption process from (6.15)

$$\begin{aligned} W_{\mathrm{ex,\,g}} &= \frac{2\pi}{\hbar} \left| \frac{-e\hbar}{\mathrm{i}m} A_0 \cdot \int \varphi_{\mathrm{ex}}^*(\boldsymbol{r}) \mathrm{e}^{\mathrm{i}(E_{\mathrm{ex}}/\hbar)t}\, \boldsymbol{e}_{\mathrm{A}}\, \mathrm{e}^{\mathrm{i}(\boldsymbol{kr} - \omega t)}\, \nabla \varphi_{\mathrm{g}}(\boldsymbol{r})\, \mathrm{e}^{-\mathrm{i}(E_{\mathrm{g}}/\hbar)t} \mathrm{d}\tau \right|^2 D(E_{\mathrm{g}} + \hbar\omega) \\ &\propto |A_0 < \varphi_{\mathrm{ex}}|\, H^{(1)}\, |\varphi_{\mathrm{g}} >|^2\, D(E_{\mathrm{g}} + \hbar\omega) \\ &=: A_0^2 \mid H_{\mathrm{ex,g}}^{(1)} \mid^2 D(E_{\mathrm{g}} + \hbar\omega)\,, \end{aligned} \tag{6.19}$$

where $\boldsymbol{e}_{\mathrm{A}}$ is the unit vector in the direction of $\boldsymbol{A}$. We also give in (6.19) a generally used abbreviation for the integral. A significant transition rate occurs only if the arguments of the time dependent exponential functions vanish, or, mathematically

$$E_{\mathrm{ex}} - E_{\mathrm{g}} - \hbar\omega = 0\,. \tag{6.20a}$$

This is again the law of energy conservation. If the φ_i and $H^{(1)}$ have plane-wave character and are described by wave vectors $\boldsymbol{k}$, a similar argument results in $\boldsymbol{k}$ conservation.

$$\hbar\boldsymbol{k}_{\mathrm{ex}} - \hbar\boldsymbol{k}_{\mathrm{g}} - \hbar\boldsymbol{k} = 0 \tag{6.20b}$$

The latter is not the case for the two-level atoms discussed here but is true for most of the eigenstates of crystalline semiconductors.

We see that the transition rate is proportional to A_0^2, and thus to the light intensity $I = \langle S \rangle$ or the density of photons $N_{\mathrm{ph}}(\omega)$ in a certain mode:

$$W_{ij} \propto A_0^2 \left| H_{\mathrm{ex,g}}^{(1)} \right|^2 \propto I \left| H_{\mathrm{ex,g}}^{(1)} \right|^2 \propto N_{\mathrm{ph}} \left| H_{\mathrm{ex,g}}^{(1)} \right|^2 \ . \tag{6.21}$$

By partial integration using the fact that the eigenfunctions form an orthonormal set, or that the $H^{(1)}$ is Hermitian adjoint, or by the argument of the microscopic reversibility of transitions from state $i \to f$ and from $f \to i$ induced by some perturbation $H^{(1)}$, we find that

$$\left| \int \varphi_f^* H^{(1)} \varphi_i \mathrm{d}\tau \right|^2 = \left| \int \varphi_i^* H^{(1)} \varphi_f \mathrm{d}\tau \right|^2 \ . \tag{6.22}$$

We see that the probabilities for induced emission and absorption are the same and that the rates differ only by factors containing the number of atoms in the upper and lower states. This fact is the basis for the relation between the so-called *Einstein coefficients* (see problem (6.1). Spontaneous emission has to be treated in the sense mentioned above as emission stimulated by the zero field. We will return to this aspect in a moment.

Now we want to simplify the interaction operator $H^{(1)}$ to reach the so-called *dipole approximation*. We note that the radius of an atom ($r \simeq 0.1\,\mathrm{nm}$) and the distance between neighboring atoms in a solid ($a \simeq 0.3\,\mathrm{nm}$) are small compared to the wavelength of light in the visible ($\lambda \simeq 500\,\mathrm{nm}$). Therefore there is practically no phase shift of the electromagnetic radiation on the length scale of one atom or of the distance between neighboring atoms. Thus we can expand the term $\mathrm{e}^{\mathrm{i}\boldsymbol{kr}}$ in (6.18) in a power series and stop after the constant term

$$\mathrm{e}^{\mathrm{i}\boldsymbol{kr}} = 1 + \frac{\mathrm{i}\boldsymbol{kr}}{1!} + \frac{(\mathrm{i}\boldsymbol{kr})^2}{2!} + \cdots \simeq 1 \ . \tag{6.23}$$

This is the first step towards the dipole approximation. It means, that the momentum of the photon $\hbar\boldsymbol{k}$ in (6.20b) is neglected.

The matrix element $H_{fi}^{(1)}$ still contains the *momentum operator* $\boldsymbol{p} = \frac{\hbar}{\mathrm{i}}\boldsymbol{\nabla}$

$$H_{fi}^{(1)} \sim \langle \varphi_f \, | \boldsymbol{p} | \, \varphi_i \rangle =: \langle \boldsymbol{p}_{fi} \rangle \ . \tag{6.24}$$

With the semiclassical relation

$$\frac{\hbar}{\mathrm{i}} \nabla = \boldsymbol{p} = m\dot{\boldsymbol{r}} \tag{6.25}$$

and some arguments of plausibility (see e.g., [55S1]), we find that

$$\int \varphi_f^* \frac{\hbar}{\mathrm{i}} \nabla \varphi_i \mathrm{d}\tau = m \frac{\mathrm{i}}{\hbar} \left(E_i - E_f\right) \int \varphi_f^* \boldsymbol{r} \varphi_i \mathrm{d}\tau = m\omega \int \varphi_f^* \boldsymbol{r} \varphi_i \mathrm{d}\tau \,. \tag{6.26}$$

For a detailed derivation of this relation actually some knowledge of the wave functions is required (for details see e.g., [55S1, 07M2]).

We note that the matrix element of the momentum operator $\boldsymbol{p} = \frac{\hbar}{\mathrm{i}} \nabla$ in the first-order perturbation term of (6.10) is directly related to the matrix element of dipole moment $e\boldsymbol{r}$. That is the reason for the naming *dipole approximation*.

The transition rate is now given by

$$W_{fi} \sim I\omega^2 \left| \boldsymbol{e}_A < e\boldsymbol{r}_{fi} > \right|^2 D(E) = I\omega^2 \left| H_{fi}^D \right|^2 D(E) \tag{6.27}$$

where $\boldsymbol{e}_A$ is the unit vector in the direction of $\boldsymbol{A}$ and $\left| H_{fi}^D \right|^2$ is just the expectation value of the dipole moment $e\boldsymbol{r}$ squared since the field amplitude squared appears already in the intensity of the light field I.

To avoid confusion we note that $\left| H_{fi}^{(1)} \right|$ contains still the field strengths $\boldsymbol{A}$ or $\boldsymbol{E}$ and has the dimension of an energy while $\left| H_{fi}^D \right|$ contains the dipole operator only.

This result can also be obtained in a more intuitive way if we remember that the energy of a dipole $e\boldsymbol{r}$ in an electric field $\boldsymbol{E} = \dot{\boldsymbol{A}}$ is given by

$$e\boldsymbol{r} \cdot \boldsymbol{E} = H^{(1)} \tag{6.28}$$

Using this approach in combination with (6.15) yields directly (6.26). Alternatively it is shown in [07M2] with similar arguments as in (6.24) that the perturbation term $-e\boldsymbol{p}\boldsymbol{A}/m$ can be replaced by $e\boldsymbol{r}\boldsymbol{E}$. Compare the perturbation term of first order in (6.10) with (6.27).

From now on we will call $\left| H_{fi}^D \right|^2$ the *dipole-transition probability* and the operator $e\boldsymbol{r}$ the *dipole operator* H^D. Transitions resulting from higher terms in (6.11), (6.17) and (6.23) correspond to quadrupole, octupole and higher-order transitions.

To conclude this section, we calculate the net rate of the transitions shown in Fig. 6.1a–c. We assume that we have a density of photons N_{ph} which populate only one mode in the sense used for the calculation of the density of states in Chap. 4 that is, all photons have the same wave vector $\boldsymbol{k}$, polarization $\boldsymbol{e}_{\mathrm{A}}$ and energy $\hbar\omega$. Furthermore $\hbar\omega$ fulfills the energy conservation law according to (6.20). The density of identical two-level atoms is N_{A} where a fraction α_{g} is in the ground state and correspondingly $(1 - \alpha_{\mathrm{g}})$ are in the excited state. The net rate of the change of N_{ph} with time is then given using ((6.21) and (6.26)) by

$$\frac{\partial N_{\mathrm{ph}}}{\partial t} = -N_{\mathrm{A}} \alpha_{\mathrm{g}} N_{\mathrm{ph}} \left| H_{\mathrm{ex\,g}}^D \right|^2 + N_{\mathrm{A}} (1 - \alpha_{\mathrm{g}})(1 + N_{\mathrm{ph}}) \left| H_{\mathrm{ex\,g}}^D \right|^2 . \tag{6.29}$$

The first term on the r.h.s. describes the absorption of photons, the second one the spontaneous and stimulated emission in the factor $(1 + N_{\mathrm{ph}})$.

From (6.22) we see that

$$\left|H^D_{\mathrm{ex,g}}\right|^2 = \left|H^D_{\mathrm{g,ex}}\right|^2 = \left|H^D\right|^2 \tag{6.30}$$

and hence

$$\frac{1}{\left|H^D\right|^2}\frac{\partial N_{\mathrm{ph}}}{\partial t} = N_{\mathrm{ph}} \cdot N_{\mathrm{A}}(1-2\alpha_{\mathrm{g}}) + N_{\mathrm{A}}(1-\alpha_{\mathrm{g}}) \ . \tag{6.31}$$

The first term on the right-hand side depends linearly on the photon density N_{ph} and describes the net rate of absorption and stimulated emission. The second terms gives the spontaneous emission since it is independent of N_{ph}.

There is net absorption for $\alpha_{\mathrm{g}} > 1/2$ (absorption coefficient $\alpha(\omega) > 0$ or $\partial N_{\mathrm{ph}}/\partial t < 0$) and amplification or optical gain for $\alpha_{\mathrm{g}} < 1/2$, i.e., for gain, more than half of the atoms have to be in the upper state. This situation cannot be reached in thermal equilibrium, but only under the influence of a suitable source of pump power. Usually one or more additional energy levels are required, i.e., lasers are typically based on *three-* or *four-level systems*.

The latter fact can be easily elucidated with the following argument. Consider a situation where all atoms are in the ground state i.e., $\alpha_{\mathrm{g}}(t=0) = 1$. If for $t > 0$ a photon field with frequency ω fulfilling (6) interacts with the system it will initially experience absorption since $\alpha(\omega)$ is given by

$$-\alpha(\omega) \sim N_{\mathrm{A}}(1-2\alpha_{\mathrm{g}}) \ . \tag{6.32}$$

With increasing time and pump power $\alpha(\omega)$ decreases because α_{g} decreases. For the situation $\alpha_{\mathrm{g}} = 1/2$ and an incoherently excited ensemble of atoms the absorption vanishes and the material becomes transparent. This means no more pumping is possible under (quasi-)stationary excitation to reach $\alpha_{\mathrm{g}} > 1/2$.

In the second volume of this book we will reexamine the above considerations for the case of strong light-matter coupling. We will see that without loss of coherence between the light field and the driven optical excitation or for sufficiently intense pulses shorter than the dephasing time it is possible to drive the system periodically from the ground state to the completely inverted state ($\alpha_{\mathrm{g}} = 0$) and back again. This process is known as Rabi flopping.

Problems

6.1. Find in a textbook the definition and meaning of Einstein's coefficients.

6.2. Inform yourself on time independent perturbation theory. The perturbed wave function contains in first order virtually excited states and the perturbed energy in second order. Inspect time dependent perturbation theory for virtual excitations. Derive Fermi's golden rule.

References

[55S1] L.I. Schiff, Quantum Mechanics, *Advanced Solid State Physics*, 2nd edn. (McGraw-Hill, New York, 1955)
[93S1] W. Stößel, *Fourieroptik* (Springer, Berlin, 1993)
[07M1] D. Meschede, *Optics, Light and Lasers*, 2nd edn. (Wiley-VCH, Weinheim, 2007)
[07M2] P. Meystre, M. Sargent III, *Elements of Quantum Optics*, 4th edn. (Springer, Berlin, 2007)
[07S1] B.E.A. Saleh, M.C. Teich, *Fundamentals of Photonics*, 2nd edn. (Wiley, Hoboken, 2007)

Chapter 7
Oscillator Model of Strong Light-Matter Coupling

Abstract The optical properties of matter are determined by the coupling of various types of "oscillators" in matter to the electromagnetic radiation field. In other words, an incident electromagnetic field will cause these oscillators to perform driven oscillations. In this *strong coupling* approach we find a typical resonance behavior where the amplitude of the driven oscillations depends on the angular frequency ω of the incident field, on the eigenfrequency ω_0 of the oscillators, on the coupling strength f between electromagnetic field and oscillator, and on its damping γ. In semiconductors the main intrinsic oscillators or *optical excitations* are *optical phonons*, *excitons* including their ionization continuum and higher band-to-band transitions or *plasmons*. Actually, we can anticipate that many basic features of the optical properties related to these oscillators are similar. Therefore it is reasonable to discuss first, in a general way, the optical properties of an ensemble of model oscillators. These model oscillators are known as *Lorentz oscillators*. By using the results of this treatment we find a quite simple and intuitive access to the optical properties of semiconductors which is in many respects very close to reality. This is obvious when comparing the calculated optical spectra and the dielectric function in the vicinity of the resonances of optical excitation to the experimental results reviewed in the remaining chapters of this book. We will thus follow this classical approach for some while, and explain at the appropriate places which modifications appear if quantum-mechanical properties are included. We begin with the simplest case of uncoupled oscillators and refine the concept in various steps in the course of this chapter.

7.1 Equations of Motion and the Dielectric Function

We consider here a model substance consisting of Lorentz oscillators and will derive its linear optical properties [63H1]. We will later generalize this concept to the *Drude–Lorentz model* [72W1, 07B1, 07S1] by including the consequences of a finite electrical conductivity $\sigma(\omega)$, e.g., in doped semiconductors. A much more rigorous theoretical treatment of semiconductor optics in terms of the *semiconductor Bloch equations* will be presented in the second volume of this book. These equations couple the polarization density P with the occupation of the semiconductor band

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_7

states. This theory enables a concise description of coherent optical phenomena, non-linear optics and dynamics.

Let us start with an ensemble of identical, uncoupled harmonic oscillators. For simplicity we choose a periodic one-dimensional array in the direction of light propagation with a lattice-constant a as shown in Fig. 7.1a. These harmonic oscillators all have the same eigenfrequency ω_0'. If we neglect damping for the moment, then ω_0' can be expressed for mechanical oscillators by the mass m and the force constant β of the springs as

$$\omega_0' = \sqrt{\beta/m}\,. \tag{7.1}$$

If we elongate the oscillators in phase (Fig. 7.1b) the whole ensemble oscillates with ω_0' and the same is true if we excite neighboring oscillators in anti-phase (Fig. 7.1c). The first case corresponds to

$$\lambda = \infty \quad \text{or} \quad k = 0\,, \tag{7.2a}$$

and the second one to

$$\lambda_{\min} = 2a \quad \text{or} \quad k_{\max} = \frac{\pi}{a}\,. \tag{7.2b}$$

The point $k = 0$ is usually called the Γ-point in $\boldsymbol{k}$-space and the condition (7.2b) gives the boundary of the first Brillouin zone for a simple linear chain or cubic lattice. The values given by (7.2b) are the shortest physically meaningful wavelength and the largest $\boldsymbol{k}$-vector in this model system.

The eigenfrequency will also be ω_0' for all λ or $\boldsymbol{k}$-values in between. Consequently, the dispersion relation for this system is horizontal (see Fig. 7.2a), i.e., the width of the band of eigenfrequencies is zero (this is called a flat band). A wave packet created in this model system by elongating only one oscillator (Fig. 7.1d) will not propagate, since there is no coupling to the neighboring oscillators. This is consistent with the fact that the group velocity $\partial\omega/\partial k$ (see (2.18)) is zero. We can generalize the findings for this example to the case of any ensemble of identical oscillators (e.g., ions in a crystalline lattice) with the following rules

$$\begin{aligned} &\text{zero coupling} \quad \longleftrightarrow \quad \text{zero bandwidth} \quad \longleftrightarrow \quad v_g = 0, \quad \text{see Fig. 7.2a} \\ &\text{finite coupling} \quad \longleftrightarrow \quad \text{finite bandwidth} \quad \longleftrightarrow \quad v_g \neq 0, \quad \text{see Fig. 7.2b}\,. \end{aligned} \tag{7.3}$$

We will discuss the implications of relaxing the assumption of the oscillators being uncoupled, proceeding to the more realistic assumption of coupled oscillators in Sect. 8.4.

In the next step we couple the independent oscillators to the electric field of a light wave given by

$$\boldsymbol{E} = (E_0, 0, 0)\exp\left[\mathrm{i}(k_z z - \omega t)\right]. \tag{7.3a}$$

This means that the light wave propagates in the z-direction and is polarized in the x-direction, parallel to the elongation of the oscillators. By considering the oscillator

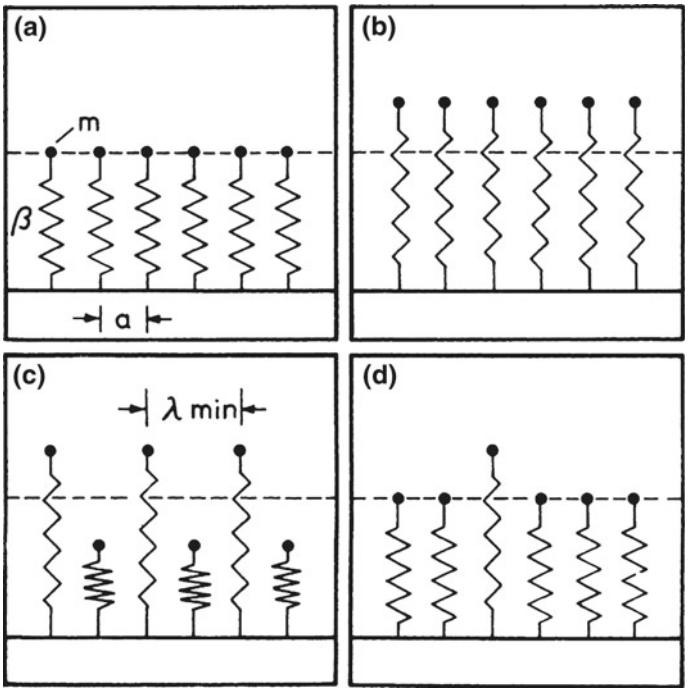


Fig. 7.1 A part of a periodic array of uncoupled oscillators in their equilibrium position (**a**), elongated with a wavelength $\lambda \rightarrow \infty$ (**b**), with the shortest physically meaningful wavelength λ_{min} (**c**), and a non-propagating wave packet (**d**) (Such depiction as mechanical oscillators was first suggested in [63H1])

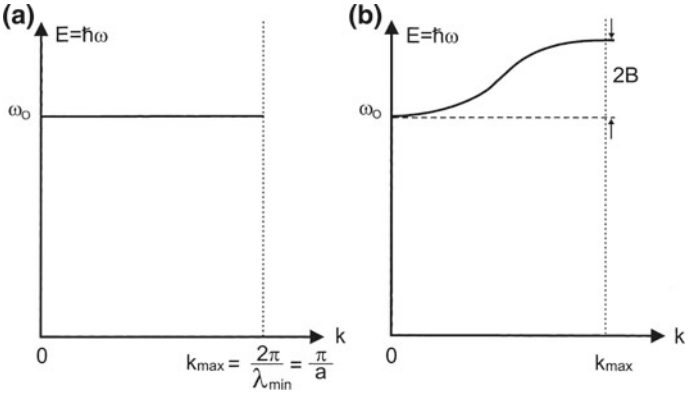


Fig. 7.2 The dispersion relation of an ensemble of uncoupled (**a**) and of coupled (**b**) oscillators

at $z = 0$ or making use of the dipole approximation of (6.23)–(6.27) (i.e., $a \ll \lambda$) we can drop the spatial variation in (7.3a) to obtain

$$\boldsymbol{E} = (E_0, 0, 0)\, \mathrm{e}^{-\mathrm{i}\omega t} . \tag{7.3b}$$

For the coupling, we assume that the mass m of every oscillator carries a charge e. For neutrality reasons, we need then a charge $-e$ fixed at the equilibrium position of every mass. The electric field $\boldsymbol{E}$ will then exert a force on the oscillator and its elongation x is connected with an electric dipole moment via

$$p_x = e\, x . \tag{7.4}$$

The oscillators described in this or in similar models are called *Lorentz oscillators* or *Selmaier oscillators*, the resulting equation for $\varepsilon(\omega)$ is also called *Helmholtz–Ketteler* or *Kramers–Heisenberg formula*.

The equation of motion of the model system coupled to light is given by

$$m\ddot{x} + \gamma m\dot{x} + \beta x = eE_0 \mathrm{e}^{-\mathrm{i}\omega t} , \tag{7.5}$$

where we have included a phenomenological damping constant γ. For a more elaborate theory, e.g., in the frame of the fluctuation–dissipation theorem or for the aspect of *Markov damping* (used here) versus non-Markovian damping see, e.g., [67K1, 85C1, 01W1, 02S1].

Equation (7.5) is a linear, inhomogeneous differential equation of second order. The general solution is a sum of the general solution of the corresponding homogeneous equation and of a special solution of the inhomogeneous one. Correspondingly we use the following ansatz:

$$x(t) = x_0 \exp\left[-\mathrm{i}\left(\omega_0'^2 - \gamma^2/4\right)^{1/2} t\right] \exp\left(-t\gamma/2\right) + x_\mathrm{p} \mathrm{e}^{-\mathrm{i}\omega t} . \tag{7.6}$$

The first term on the right-hand side is the solution of the homogeneous equation and describes a transient feature. For

$$\omega_0'^2 - \gamma^2/4 > 0 \tag{7.7}$$

one finds a damped oscillation with a damping-dependent eigenfrequency $\left(\omega_0'^2 - \gamma^2/4\right)^{1/2}$. The inequality (7.7) defines the regime of *weak damping*. For stronger damping one gets essentially an exponentially decaying term.

This transient feature disappears in any case after $t > \gamma^{-1}$. It is thus of importance for (ultra-)fast dynamics treated in the second volume of this book. For the stationary, linear optics regime, in which we are presently interested, we may safely omit this term. What is then left is a forced oscillation with amplitude x_p. Inserting the ansatz (7.6) into (7.5) we find

$$x_\mathrm{p} = \frac{eE_0/m}{\left(\omega_0'^2 - \omega^2 - \mathrm{i}\omega\gamma\right)}\,, \tag{7.8}$$

i.e., the usual resonance term of a driven oscillation.

This oscillation is connected with a dipole moment at every oscillator of

$$p_x = e\,x_\mathrm{p} \tag{7.9}$$

and a *polarizability* $\hat{a}(\omega)$:

$$\hat{a}(\omega) = \frac{ex_\mathrm{p}}{E_0} = \frac{e^2/m}{\left(\omega_0'^2 - \omega^2 - \mathrm{i}\omega\gamma\right)}\,. \tag{7.10}$$

The hat on $\hat{a}$ and the prime on ω_0' indicate that we need to introduce two corrections to the here introduced quantities which will be done in Sect. 7.2.

If we use a three-dimensional array of oscillators with density N, the result is the same and we shall get a preliminary *polarization density* $\boldsymbol{P}$ with amplitude $\boldsymbol{P}_0$

$$\boldsymbol{P}_0 = N\hat{a}\boldsymbol{E}_0 = \frac{Ne^2/m}{\left(\omega_0'^2 - \omega^2 - \mathrm{i}\omega\gamma\right)}\boldsymbol{E}_0\,. \tag{7.11}$$

This means that we describe the ensemble of oscillators as an *effective medium*, an approach which is well justified for $a \ll \lambda$.

From (7.11) to (2.30) we get the following expressions for the *dielectric displacement* $\boldsymbol{D}$, the *dielectric function* $\varepsilon(\omega)$, and the *susceptibility* $\chi(\omega)$:

$$\boldsymbol{D} = \varepsilon_0\boldsymbol{E} + \boldsymbol{P} = \varepsilon_0\left[1 + \frac{Ne^2/\varepsilon_0 m}{\left(\omega_0'^2 - \omega^2 - \mathrm{i}\omega\gamma\right)}\right]\boldsymbol{E} \tag{7.12}$$

and

$$\varepsilon(\omega) = 1 + \frac{Ne^2/\varepsilon_0 m}{\left(\omega_0'^2 - \omega^2 - \mathrm{i}\omega\gamma\right)} = \chi(\omega) + 1\,. \tag{7.13}$$

In the following we want to address two corrections to the above set of equations in order to obtain the final expression for $\varepsilon(\omega)$.

7.2 Corrections Due to Quantum Mechanics and Local Fields

The term $Ne^2/\varepsilon_0 m$, e.g., in (7.11) gives the coupling strength of the electromagnetic field to the oscillators in the mechanical model system. In quantum mechanics, this coupling is given by the transition matrix element squared. For dipole-allowed

transitions this reads, as mentioned already in Sect. 6.2:

$$\left|H_{ij}^{D}\right|^2 = \left|\langle j \left| H^{D} \right| i \rangle\right|^2 , \tag{7.14}$$

where $|H_{ij}^{D}|^2$ gives the expectation value of the *dipole operator* squared. Here i and j stand for initial and final state and H^D for the dipole operator $\boldsymbol{er}$ or ex (see also the comments related to (6.27) and (6.28)). For dipole-forbidden transitions, magnetic dipole or electric quadrupole matrix elements can become relevant. They are usually orders of magnitude smaller.

There are different conventions for introducing the transition matrix element into the dielectric function. Some authors define a dimensionless quantity $\hat{f} = 2m\omega_0' \left|H_{ij}^{D}\right|^2 / (\hbar e^2)$, call it the *oscillator strength*, and use it to multiply the term $Ne^2m^{-1}\varepsilon_0^{-1}$. Others call the whole numerator oscillator strength f, i.e.,

$$f = \frac{2N\omega_0'}{\varepsilon_0 \hbar} \left|H_{ij}^{D}\right|^2 . \tag{7.15}$$

We follow the second approach for simplicity of notation, but stress that there is no physical difference between the definitions. We obtain

$$\varepsilon(\omega) = 1 + \frac{f}{\omega_0'^2 - \omega^2 - \mathrm{i}\omega\gamma} . \tag{7.16}$$

Sometimes the numerator and denominator on the r.h.s of (7.16) are multiplied by $\hbar^2$. Then the oscillation strength has the dimension of an energy squared, e.g., $(\mathrm{eV})^2$.

For the next correction to (7.16) we have to consider what is the actual electric field $\boldsymbol{E}$ (e.g., in (7.5)) acting on the oscillators. For dilute systems, $\boldsymbol{E}$ is just the external incident field and we can use (7.16) as it is. For dense systems, i.e., systems with a high density N of oscillators, the *local field* E_{loc} acting on the oscillators consists of two parts, namely the external field and the field created by all the other dipoles of the dense medium. Taking into account this effect leads for cubic or amorphous materials to the so-called *Clausius–Mosotti* or *Lorenz–Lorentz equation* (see also [02K1, 07B1]), which relates the polarizability $\hat{a}$ to $\varepsilon(\omega)$ through

$$\frac{\varepsilon(\omega) - 1}{\varepsilon(\omega) + 2} = \frac{N\hat{a}(\omega)}{3\varepsilon_0} = \frac{\frac{1}{3} N \frac{e^2}{m\varepsilon_0}}{\omega_0'^2 - \omega^2 - \mathrm{i}\omega\gamma} . \tag{7.17a}$$

Obviously (7.17a) recovers the form (7.16) for dilute systems (small N) for which $N\hat{a}$ is small and $\varepsilon(\omega)$ deviates only a little from unity. If this approximation is not valid, $\varepsilon(\omega)$ can be rewritten for small damping in the form (7.16) but with a shifted eigenfrequency

$$\omega_0^2 = \omega_0'^2 - \frac{Ne^2}{3m\varepsilon_0} = \omega_0'^2 - f/3 . \tag{7.17b}$$

resulting in

$$\varepsilon(\omega) = 1 + \frac{f}{\omega_0^2 - \omega^2 - \mathrm{i}\omega\gamma} \,. \tag{7.18}$$

This formula now incorporates local field effects and the quantum mechanical transition probabilities. The new eigenfrequency ω_0 is the only physically relevant and experimentally accessible one. A similar procedure is also valid for crystal symmetries other than the cubic one, but in these cases the tensor character of $\varepsilon(\omega)$ comes into play, resulting in parameters f, ω_0 and γ depending on the orientation of the electric field relative to the crystal axes.

7.3 Spectra of the Dielectric Function and of the Complex Index of Refraction

As we shall see later, a semiconductor contains not only one type of oscillators and one resonance frequency ω_0, but many of them—like phonons, excitons etc. In linear optics, i.e., in linear response theory, we can just sum over all resonances leading to

$$\varepsilon(\omega) = 1 + \sum_j \frac{f_j}{\omega_{0j}^2 - \omega^2 - \mathrm{i}\omega\gamma_j} \,. \tag{7.19a}$$

This is essentially the so-called *Helmholtz–Ketteler formula* or *Kramers–Heisenberg dielectric function*.[1]

If the eigenfrequencies ω_0 form a continuous band, like in the ionisation continuum of the excitons or in an inhomogeneously broadened system where every oscillator has a slightly different frequency ω_0, then the Σ in (7.19a) has to be replaced by an integral over these contributions

$$\int \frac{f(\omega_0)}{\omega_0^2 - \omega^2 - \mathrm{i}\omega\gamma(\omega_0)} \mathrm{d}\omega_0 \tag{7.19b}$$

where $f(\omega_0)$ is the *oscillator strength* per frequency interval.

It should be noted, that the case of several resonances in the dielectric function cannot only be written as the sum of (7.19a), which is a logic consequence of the approach of linear optics, but also as a product (see [74G1, 06A1]):

$$\varepsilon(\omega) = \prod_j \frac{\omega^2 - \omega_{\mathrm{L0},j} + \mathrm{i}\gamma_{\mathrm{L0},j}\omega}{\omega^2 - \omega_{\mathrm{T0},j} + \mathrm{i}\gamma_{\mathrm{T0},j}\omega} \,. \tag{7.19c}$$

[1] Sometimes the oscillator strength is defined in a slightly different way as F_j with $f_j = F_j\omega_{0j}^2$.

This formula has the advantage, that different damping values γ can be introduced for the transverse (TO) and longitudinal (LO) states of the same resonance. If they are equal (7.19c) can be rewritten in the form (7.19a) to (7.19b). For simplicity of writing we will use the form (7.19a) in the following.

We now want to discuss the contribution of an *isolated resonance* at $\omega_{0j'}$ in (7.19a) which allows some simplification. For closely spaced resonances the whole formula (7.19a) has to be used. We discuss this situation separately in Sect. 7.5. For the present purpose of a single resonance well separated from all other resonances, we note that the contribution of this single resonance at $\omega_{0j'}$ to the dielectric function $\varepsilon(\omega)$ in (7.19a) is constant for light frequencies $\omega \ll \omega_{0j'}$ with a value of

$$\frac{f_{j'}}{\omega_{0j'}^2} = \text{const} \tag{7.20}$$

and tends to zero for $\omega \gg \omega_{0j'}$.

This means that in the vicinity of an isolated resonance $\omega_{0j'}$ we can neglect the contributions from all lower resonances $\omega_{0j} \ll \omega_{0j'}$ and the constant contributions of all higher resonances $\omega_{0j} \gg \omega_{0j'}$ can be summarized in a so-called *background dielectric constant* ε_b.[2] Obviously ε_b is unity for the highest resonance in the system, i.e. in the sum of (7.19a). For our purposes this highest resonance is typically in the X-ray regime connected, e.g., with the K absorption edge. So we finally get the simplified expression in the spectral surroundings of an isolated resonance at $\omega_{0j'}$:

$$\varepsilon(\omega) = \varepsilon_\mathrm{b} + \frac{f_{j'}}{\omega_{0j'}^2 - \omega^2 - \mathrm{i}\omega\gamma_{j'}} \tag{7.21a}$$

In the following we drop the index j'.

Equation (7.21a) can be separated into real and imaginary parts

$$\begin{aligned}\varepsilon(\omega) &= \varepsilon_\mathrm{b} + \frac{f(\omega_0^2 - \omega^2)}{(\omega_0^2 - \omega^2)^2 + \omega^2\gamma^2} + \mathrm{i}\frac{\omega\gamma f}{(\omega_0^2 - \omega^2)^2 + \omega^2\gamma^2} \\ &= \varepsilon_1(\omega) + \mathrm{i}\varepsilon_2(\omega)\ .\end{aligned} \tag{7.21b}$$

For vanishing damping $\gamma \to 0$ and in the vicinity of a resonance ω_0 equation (7.21a) can be simplified by noting $(\omega_0^2 - \omega^2) = (\omega_0 + \omega)(\omega_0 - \omega) \approx 2\omega_0(\omega_0 - \omega)$ to

$$\varepsilon(\omega) = \varepsilon_\mathrm{b} + \frac{f}{2\omega_0}\frac{1}{\omega_0 - \omega} \tag{7.21c}$$

The resonance behavior of the real and imaginary parts of the *dielectric function* $\varepsilon(\omega)$ are shown in Fig. 7.3. For negligible *damping* $\gamma \to 0$ we find a pole in $\mathrm{Re}\{\varepsilon(\omega)\}$, and $\mathrm{Im}\{\varepsilon(\omega)\}$ converges to a δ–function at ω_0. Finite damping results in a broadening

[2] In the notation of (7.19c) ε_b appears as a prefactor to the product instead of a summand in (7.21a).

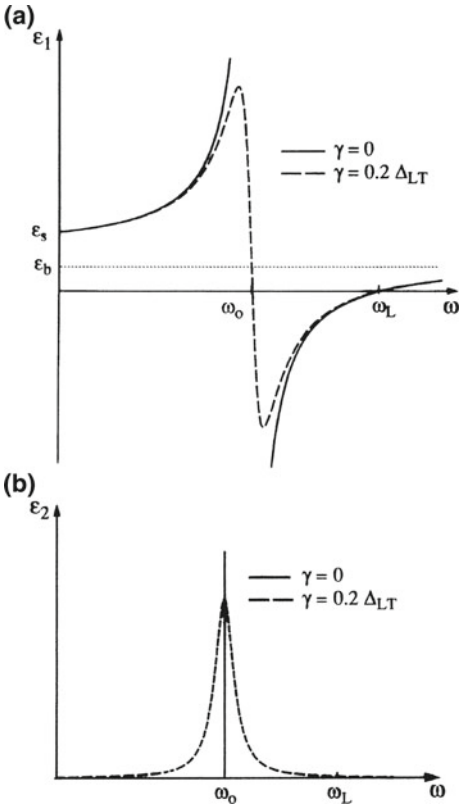


Fig. 7.3 The real (**a**) and imaginary (**b**) parts of the dielectric function for zero and finite damping

of $\mathrm{Im}\{\varepsilon(\omega)\}$ to the *Lorentzian lineshape* of (7.21b) and a smooth connection of the two branches of $\mathrm{Re}\{\varepsilon(\omega)\}$.

We concentrate now on the case of small damping. One of the two frequencies of special interest for $\varepsilon(\omega)$ is the eigenfrequency ω_0, which is connected with the singularity. The other one corresponds to the point at which $\mathrm{Re}\{\varepsilon(\omega)\}$ crosses zero. Going back to (2.20) and (2.46) we find

$$\nabla \cdot \boldsymbol{D} = \varepsilon_0 \nabla(\varepsilon(\omega) \cdot \boldsymbol{E}) = 0 \,. \tag{7.22}$$

As already discussed, this equation is usually used to argue that electromagnetic waves are transverse, since $\nabla \cdot \boldsymbol{E} = 0$ for this case. The other solution $\varepsilon(\omega = \omega_\mathrm{L}) = 0$ gives the frequency of a longitudinal mode. This mode is a pure polarization mode with $\varepsilon_0 \boldsymbol{E} = -\boldsymbol{P}$, i.e., anti-parallel polarization density and electric field (see Sect. 2.3).

For $\gamma \to 0$ we find

$$\begin{aligned} \varepsilon(\omega = \omega_{\mathrm{T}} \equiv \omega_0) &= \infty, \\ \varepsilon(\omega = \omega_{\mathrm{L}}) &= 0. \end{aligned} \tag{7.23}$$

The relation between these two frequencies (called *transverse* and *longitudinal eigenfrequency*, respectively) is given by

$$\omega_{\mathrm{L}}^2 - \omega_{\mathrm{T}}^2 = f/\varepsilon_{\mathrm{b}} \sim \left| H_{ij}^D \right|^2 . \tag{7.24a}$$

In the approximation (7.21c) the relation (7.24a) reduces to

$$\omega_{\mathrm{L}} - \omega_{\mathrm{T}} = f/(2\omega_{\mathrm{T}}\varepsilon_{\mathrm{b}}). \tag{7.24b}$$

We can summarize the current discussion of the resonance behavior of the dielectric function $\varepsilon(\omega)$ of our model system by the following statements. A finite oscillator strength f (i.e., a finite light-matter coupling) is necessarily connected with a finite *longitudinal-transverse splitting energy* Δ_{LT} and vice versa:

$$\left| H_{ij}^D \right|^2 \neq 0 \Leftrightarrow f \neq 0 \Leftrightarrow \Delta_{\mathrm{LT}} = \hbar(\omega_{\mathrm{L}} - \omega_{\mathrm{T}}) \neq 0 \, . \tag{7.25}$$

The physical reason for this energy splitting is that the longitudinal polarization wave produces, in contrast to the transverse wave a longitudinal electric field (see discussion following (2.46)), which acts as an additional restoring force and increases the longitudinal eigenfrequency ω_{L} above the transverse eigenfrequency. This is illustrated in Fig. 7.4 for ions in a zincblende lattice. An intuitive explanation given in [10Y1] argues that the charged ion planes act like capacitor plates. Separating them in a longitudinal wave leads to additional forces. In the case of a transverse wave the capacitor plates are moving parallel to each other which does not lead to additional restoring forces.

An important formula which relates the characteristic frequencies of the resonance to the change of dielectric function resulting from the resonance is the Lyddane–Sachs–Teller relation. We arrive at this relation if we call the roughly constant value of $\varepsilon(\omega)$ well below ω_0 the *static dielectric constant* ε_{s} and the value above ε_{b}. These quantities are related to the oscillator strength f by

$$\varepsilon_{\mathrm{s}} = \varepsilon_{\mathrm{b}} + f/\omega_0 \, . \tag{7.26}$$

The *Lyddane–Sachs–Teller relation* now reads:

$$\frac{\varepsilon_{\mathrm{s}}}{\varepsilon_{\mathrm{b}}} = \frac{\omega_{\mathrm{L}}^2}{\omega_0^2} > 1 \quad \text{for} \quad f > 0 \, . \tag{7.27}$$

Furthermore we can state that for well isolate resonances the background dielectric constant ε_{b} of a resonance ω_{0j} is simultaneously the static dielectric constant for the next higher resonance ω_{0j+1}.

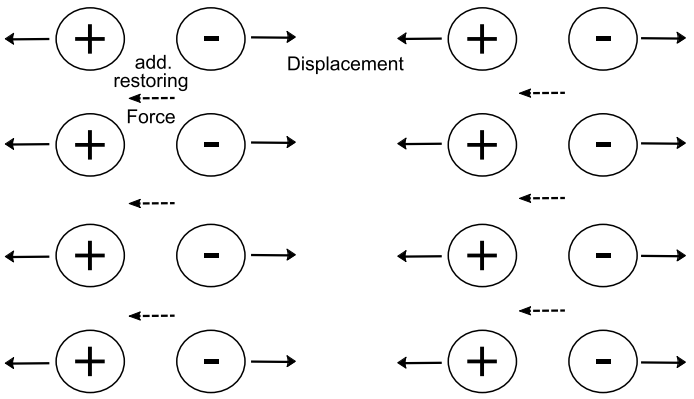


Fig. 7.4 The displacement of ions for a longitudinal phonon mode moving from left to right along a < 110 > direction in a zincblende semiconductor and the resulting additional restoring forces (adapted by permission from [10Y1] ©2010 Springer Nature)

Finally we can define the term *small damping* already used several times above by

$$\gamma < \hbar^{-1}\Delta_{\mathrm{LT}} = \omega_{\mathrm{L}} - \omega_{\mathrm{T}} \tag{7.28}$$

and by frequencies ω far above or below ω_0 we mean

$$\hbar\,|\omega - \omega_0| \gg \Delta_{\mathrm{LT}}\,. \tag{7.29}$$

To describe the *propagation of light* in the model oscillator system we will now concentrate on the *complex index of refraction* $\tilde{n}(\omega) = n(\omega) + \mathrm{i}\kappa(\omega)$, which is connected to the dielectric function via (2.38)

$$\tilde{n}(\omega) = \varepsilon^{1/2}(\omega)\,. \tag{7.30}$$

This leads to the following relations between the various real and imaginary parts:

$$\begin{aligned}\varepsilon_1(\omega) &= n^2(\omega) - \kappa^2(\omega)\\ \varepsilon_2(\omega) &= 2n(\omega)\kappa(\omega)\end{aligned}$$

or

$$\begin{aligned}n(\omega) &= \left(\frac{1}{2}\left\{\left[\varepsilon_1^2(\omega) + \varepsilon_2^2(\omega) + \varepsilon_1(\omega)\right]^{1/2}\right\}\right)^{1/2}\\ \kappa(\omega) &= \left(\frac{1}{2}\left\{\left[\varepsilon_1^2(\omega) + \varepsilon_2^2(\omega) - \varepsilon_1(\omega)\right]^{1/2}\right\}\right)^{1/2}\,.\end{aligned} \tag{7.31}$$

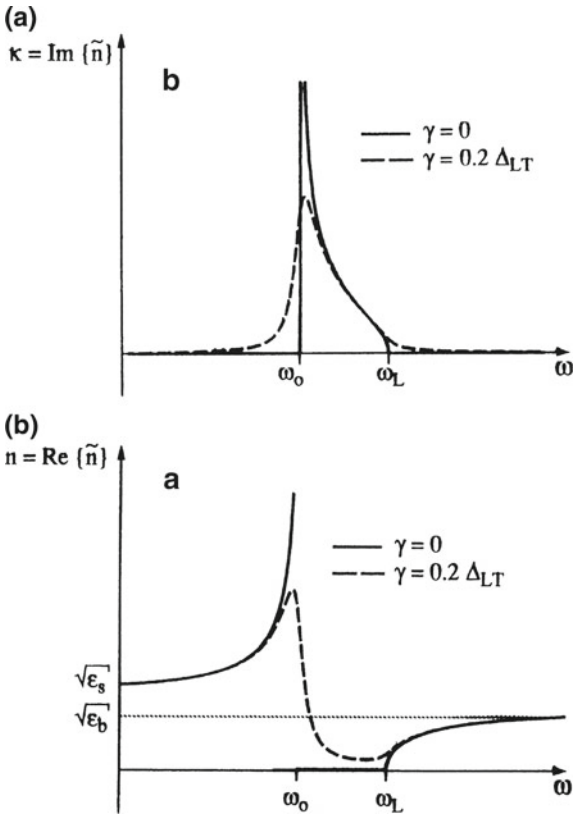


Fig. 7.5 Real (**a**) and imaginary (**b**) parts of the complex index of refraction for zero and finite damping

These real and imaginary parts of $\tilde{n}$ are illustrated in Fig. 7.5 for vanishing and for finite damping γ.

When approaching the resonance from the low-frequency side, one sees that both $\mathrm{Re}\{\tilde{n}\} = n$ and thus also the real part of $k = \frac{\omega}{c} n(\omega)$ increase drastically. Between ω_0 and ω_L we find

$$\left.\begin{array}{lll} n = 0 & \text{for} & \gamma = 0 \\ n \ll 1 & \text{for} & \gamma \neq 0 \\ \kappa > n & \text{for} & \text{all}\,\gamma \end{array}\right\} \quad \text{for} \quad \omega_0 \leq \omega \leq \omega_\mathrm{L}\,. \tag{7.32}$$

The imaginary part of $\tilde{n}(\omega)$, i.e., $\kappa(\omega)$ starts with a singularity at $\omega = \omega_0$ for $\gamma = 0$ and then drops to small values for ω approaching ω_L. This means that for $\gamma = 0$ there is no propagating, i.e., no spatially oscillating or wave-like solution in the medium for the spectral region addressed in (7.32). Instead there is only an amplitude decaying exponentially in space similar to the type known for total reflection in the medium with the lower index of refraction. For finite γ there is a small real part of $\tilde{n}$ which means that some light can penetrate into the medium. But, this light is damped over

a distance shorter than the wavelength in the medium since $\kappa > n$. This is called the regime of *strong absorption*.

If all oscillators have the same eigenfrequency ω_0 and the broadening in $\varepsilon_2(\omega)$ results only from a finite damping γ, the system is said to be *homogenously broadened* and $\varepsilon_2(\omega)$ shows the usual Lorentzian line shape. If the system contains many oscillators with slightly different frequencies ω_0, the system is said to be *inhomogeneously broadened*. The shape of $\varepsilon_1(\omega)$ and $\varepsilon_2(\omega)$ is then determined by the distribution function of eigenfrequencies and/or of oscillator strengths in (7.19b). The width of $\varepsilon_2(\omega)$ is then a convolution of a Lorentzian $L(\omega)$ for every single oscillator and of the distribution. In the case of a Gaussian $G(\omega)$ this convolution leads to a *Voigt lineshape* in the sense of

$$V(\omega) = \int_{-\infty}^{+\infty} L(\omega - \omega')G(\omega')\mathrm{d}\omega' = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{+\infty} \frac{\Gamma}{\sigma} \frac{\exp\left\{-\dfrac{(\omega' - \omega_0)^2}{2\sigma^2}\right\}}{(\omega - \omega')^2 + \left(\dfrac{\Gamma}{2}\right)^2} \mathrm{d}\omega' \tag{7.33}$$

where σ and Γ give the inhomogeneous and homogenous broadening, respectively.

It should also be noted that even for homogeneously broadened resonances $\kappa(\omega)$ and the absorption coefficient $\alpha(\omega)$ do not exhibit intrinsically a Lorentzian lineshape in contrast to $\varepsilon_2(\omega)$ (see, e.g., Fig. 7.4b or (7.31). However for small oscillator strength and/or concentration of the oscillators and finite damping, $\alpha(\omega)$ can recover a Lorentz-like lineshape (see, e.g., [76P1]).

7.4 The Spectra of Reflection and Transmission

With knowledge of $\tilde{n}(\omega)$ we can now deduce the spectra of reflection and transmission of the model substance. We start with the *reflectivity* $R(\omega)$ of a single interface between vacuum (or air) and the medium. We discuss only the situation of normal incidence (see Sect. 5.1.4 and (5.19)).

$$R(\omega) = \frac{I_\mathrm{r}}{I_\mathrm{i}} = \frac{[n(\omega) - 1]^2 + \kappa^2(\omega)}{[n(\omega) + 1]^2 + \kappa^2(\omega)} \tag{7.34a}$$

Oblique incidence is described by Fresnel's formulae (5.18).

The reflectivity under normal incidence $R(\omega)$ is plotted in Fig. 7.6. First we discuss the situation for vanishing damping $\gamma = 0$. When approaching the resonance from the low frequency side, R starts with an almost constant value determined from (5.19b), (7.26), and (7.27):

$$R = \frac{\left(\varepsilon_\mathrm{s}^{1/2} - 1\right)^2}{\left(\varepsilon_\mathrm{s}^{1/2} + 1\right)^2} \quad \text{for} \quad \omega_0 - \omega \gg \frac{1}{\hbar}\Delta_\mathrm{LT}\,. \tag{7.34b}$$

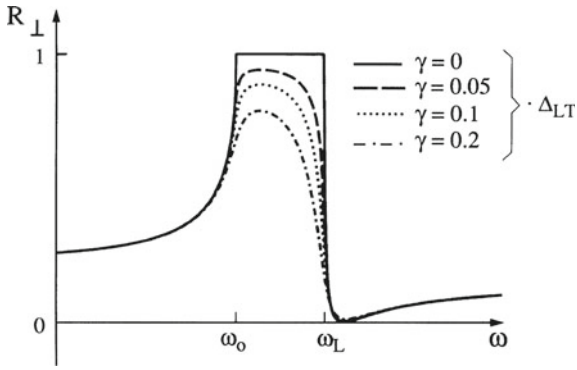


Fig. 7.6 The reflection spectrum of a single resonance with zero and finite damping for normal incidence

Above the longitudinal eigenfrequency ω_L, the reflectivity converges towards a lower constant value

$$R = \frac{\left(\varepsilon_b^{1/2} - 1\right)^2}{\left(\varepsilon_b^{1/2} + 1\right)^2} \quad \text{for} \quad \omega - \omega_0 \gg \frac{1}{\hbar}\Delta_{LT} \,. \tag{7.34c}$$

For vanishing damping, the reflectivity R increases just below ω_0 and reaches the value $R = 1$ for $\omega = \omega_0$. Between ω_0 and ω_L the reflectivity is unity, i.e., $R = 1$. For $\omega > \omega_L$ R drops rapidly and reaches zero at the frequency where $n(\omega) = 1$. Then R increases again towards a constant value given by (7.34c). It is obvious that all light is reflected in the region where no propagating mode exists in the medium in agreement with the discussion of (7.32). Such a band of high reflectivity is also called *stop band* because the light is stopped and sent back.

An alternative approach to understanding the stop band is via total internal reflection (see (5.18) or Fig. 5.6b). For vacuum the refractive index is $n \equiv 1$. Between ω_0 and ω_L the value of n is below unity. Consequently, matter is here optically "thinner" than a vacuum. In such a situation, total "internal" reflection occurs for angles of incidence exceeding the critical angle α_c. In general, α_c is defined such that $\sin\alpha_c$ cannot be larger than unity resulting for (5.11a) in the condition

$$\sin\alpha_c = \frac{n_{II}}{n_I} \quad \text{for} \quad n_{II} < n_I \tag{7.35}$$

where $n_I \equiv 1$. For $n_{II} \to 0$ we find $\alpha_c \to 0$ resulting in total internal reflection already for normal incidence.

For finite damping γ the reflection spectrum is smoothed out due to the finite values of $n(\omega)$ for $\omega_0 \le \omega \le \omega_L$ and of $\kappa(\omega)$ for $\omega \le \omega_0$ and for $\omega_L \le \omega$. It should be noted, that the reflection minimum occurs for weak damping still slightly above ω_L, but the reflection maximum is no longer related to ω_0. We shall see in Sect. 8.4 that this is even true for $\gamma = 0$ in the case of resonances with spatial dispersion.

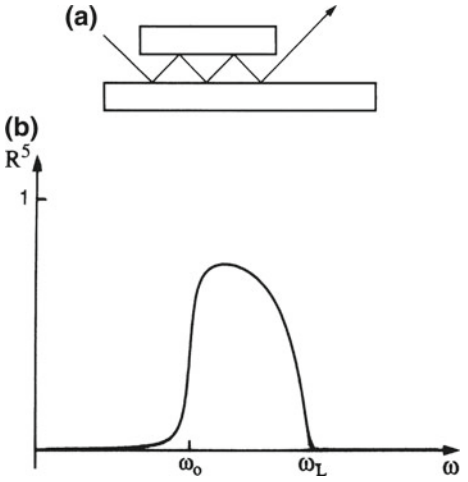


Fig. 7.7 An arrangement for multiple reflection from a medium (**a**) and the resulting Reststrahlen band (**b**)

The spectral region between ω_0 and ω_L is often called *Reststrahlen band*[3] for the following reason. If we send a light beam with a broad, essentially flat, spectral distribution of the intensity around ω_0 onto a sample and allow it to be reflected several times, e.g., in the configuration given in Fig. 7.7a, then significant intensity will remain as a "rest" only in the region $\omega_0 \leq \omega \leq \omega_L$. It should be noted that we do not assume any coherent superposition of the beams, i.e., the Reststrahlen band comes only from $\varepsilon(\omega)$ and has nothing to do with a Fabry–Perot resonator (Sect. 5.3).

Up to now we have discussed the reflectivity of a single interface. In semiconductor spectroscopy or in optical characterization of semiconductor devices containing thin films one typically has to consider *plane-parallel slabs* or films of matter with two surfaces and a geometrical thickness d rather than samples filling a semi-infinite half-space. To handle this problem, we use the results of Sect. 5.3.

In the vicinity of the resonance the reflection spectrum will remain the same as in Fig. 7.6 since multiple reflection is suppressed due to the large value of κ and thus of the absorption coefficient (2.43)

$$\alpha(\omega) = 2\,\frac{\omega}{c}\,\kappa(\omega). \tag{7.36}$$

Away from the resonance we have to consider for the total *transmission* $\hat{T}$ contributions due to multiple reflections at the two surfaces. There are two cases, as already discussed in Sect. 5.3. In the first case d is larger than the coherence length of the light source, or the two surfaces of the slab are not exactly plane-parallel (e.g., due to steps on the surfaces or due to a small tilt angle). In this case we have to add

[3]This term has been introduced in 1887 by H. Rubens and is used also in the English literature as a synonym for stop band.

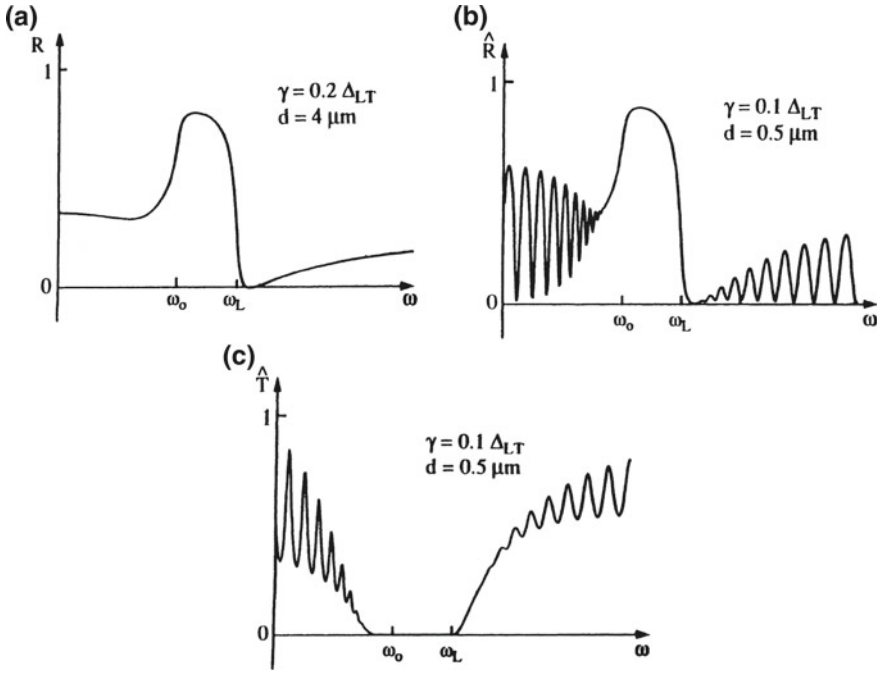


Fig. 7.8 The reflectivity $\hat{R}$ of a thin slab of matter in the vicinity of a resonance for the cases where intensities (**a**) or field amplitudes (**b**) have to be added and the transmission (**c**) for the case where amplitudes add

intensities. This results in a total transmission $\hat{T}$ and reflection $\hat{R}$ of the slab given by (5.29).

In Fig. 7.8a we show $\hat{R}$ for this situation. Below the resonance region the reflectivity is given by the combined contributions of front and rear side. Approaching the resonance, $\hat{R}$ drops first slightly because the absorption reduces the influence of the rear side. The reflection around the resonance remains the same as above in Fig. 7.6. If the sample becomes transparent above the resonance, we again find a contribution from the rear side to the reflectivity.

In Figs. 7.8b,c we give an example for the $\hat{T}$ and $\hat{R}$ spectra for the second case when the sample forms a Fabry–Perot (FP) resonator using the set of equations (5.31), and (5.32). We see from the modulation depth that the finesse F is reduced close to the resonance due to increasing absorption. The spectral spacing between the FP resonances decreases as the resonance region is approached from lower frequencies, due to the steep increase of $n(\omega)$ or of $\boldsymbol{k}$. For thin slabs of matter the spectral position of the last visible FP mode(s) below a resonance may come spectrally close to the resonance and the spacing, width and height of this last visible FP mode may be comparable to those of the resonance itself. This fact may lead to misinterpretations as discussed in [07K1, 10K1] or in Chap. 6 of [10K2]. The FP modes may reappear above the resonance, if the sample becomes transparent again.

Every directional dependence of the quantities ω_0, f and γ which describe a resonance will lead to a directional dependence of $n(\omega)$ and $\kappa(\omega)$, i.e., to birefringence and dichroism as discussed in Sect. 5.4.1. In Sect. 8.4 we shall see that a dependence of the dispersion $\omega_0(\boldsymbol{k})$ can lead to a weak observable birefringence. For a cubic crystal with point group T_d (see Appendix B) one finds that light propagating with $\boldsymbol{k}$ parallel to the [110] direction may display slightly different optical properties when the electric field is polarized in the directions [110] and [001].

Light that has been absorbed in matter is often re-emitted as luminescence light. At this point we do not consider such spectra of luminescence since they require more detailed knowledge of the excitations in the semiconductor and especially of the radiative and nonradiative decay channels. Examples will be given in Chaps. 22 and 23.

7.5 Interaction of Close-Lying Resonances

Until now we have considered a single resonance, well separated from others. Now we discuss what happens for *two close-lying resonances*. Such a situation occurs, for instance, for the A and B Γ_5 excitons in ZnO (see e.g., [77L1, 81K1, 93F1, 02K1]. To this end we consider two resonances with equal oscillator strength f and slightly different eigenfrequencies $\omega_\mathrm{A} < \omega_\mathrm{B}$. For simplicity we neglect damping and spatial dispersion. We will introduce the latter only in Sect. 8.4, but it would actually be adequate for exciton resonances due to their $\boldsymbol{k}$-dependent resonance frequency $\omega_0(\boldsymbol{k})$.

The equations displayed graphically in Fig. 7.9 are:

$$\varepsilon_\mathrm{A}(\hbar\omega) = \varepsilon_\mathrm{b} + \frac{f}{\hbar^2\left(\omega_\mathrm{A}^2 - \omega^2\right)} \tag{7.37a}$$

$$\varepsilon_\mathrm{B}(\hbar\omega) = \varepsilon_\mathrm{b} + \frac{f}{\hbar^2\left(\omega_\mathrm{B}^2 - \omega^2\right)} \tag{7.37b}$$

$$\begin{aligned}\varepsilon_\mathrm{tot}(\hbar\omega) &= \varepsilon_\mathrm{b} + \frac{f}{\hbar^2\left(\omega_\mathrm{A}^2 - \omega^2\right)} + \frac{f}{\hbar^2\left(\omega_\mathrm{B}^2 - \omega^2\right)} \\ &= \varepsilon_\mathrm{A}(\hbar\omega) + \varepsilon_\mathrm{B}(\hbar\omega) - \varepsilon_\mathrm{b}\,. \end{aligned} \tag{7.37c}$$

The graphs in Fig. 7.9a,b are the expected ones considering Fig. 7.3a. The longitudinal transverse splitting Δ_LT (about 5 meV, in each case) is the same for both resonances, because we used the same oscillator strength f and background dielectric constant ε_b for both resonances.

The situation becomes significantly different for (7.37c) shown in Fig. 7.9c. Though the oscillator strengths are still the same for both resonances, the value of Δ_LT is now smaller for resonance A and larger for resonance B, as indicated by their horizontal "shifts" on the energy axis while the positions of the singularities ω_A

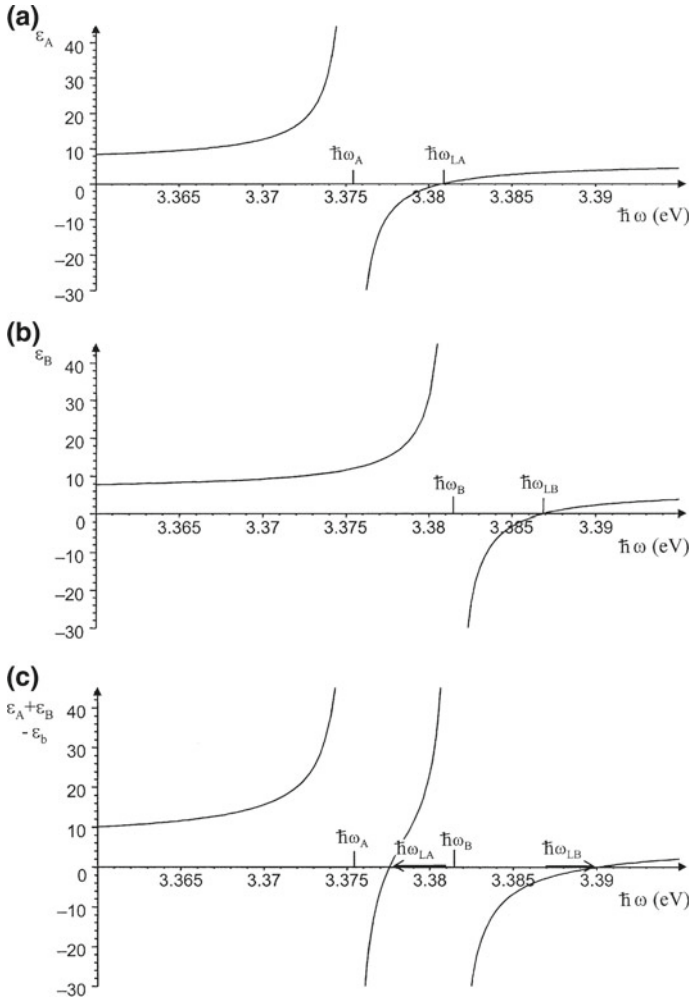


Fig. 7.9 The real part of $\varepsilon(\hbar\omega)$ for two resonances separately (**a**, **b**) and for two close lying resonances (**c**). Values on the x-axis refer roughly to the excitonic resonances in ZnO

and ω_B did not shift. The splitting Δ_{LT} decreased to about 2 meV for the resonance A and increased to about 8 meV for resonance B. This is a first example of *level repulsion*, which we will encounter more frequently in the following.

We finally want to compare the discussed scenario of excitons in ZnO to the one of the AΓ_5 and BΓ_5 exciton resonances of CdS. The latter are actually sufficiently separated in energy (several times Δ_{LT}) to treat them to a good approximation as independent resonances, in contrast to case of ZnO. We will come back to a detailed discussion in Sect. 21.1 (see also [77L1]).

The moral of this section is, that in general one cannot deduce from Δ_{LT} or the spectral width of the stop band the oscillator strengths of two close-lying resonances in a straight-forward way [01G1]. One rather has to perform a careful analysis.

Problems

7.1. Study for the case of weak damping some reflection spectra in the infrared (optical phonons) of at least partly ionic bound semiconductors, and compare with the data for ω_0 and ω_L given there.

7.2. Calculate the spectra of reflection for a single surface for weak and for strong damping with otherwise constant parameters. Compare the shift of the reflection maxima and minima with respect to the transverse and longitudinal eigenenergies ω_0 and ω_L, respectively. Which quantity can be deduced with reasonable accuracy from a first inspection of the reflection spectra?

7.3. Make a qualitative sketch of the electric fields for normal incidence of light on a medium with higher or lower index of refraction and weak or vanishing absorption, for strong absorption and for frequency $\omega_0 < \omega < \omega_L$ and vanishing damping.

7.4. Compare the experimental preparation of the Reststrahlen band in Fig. 7.7a to the functional principle of a so-called Lummer–Gehrcke device. The latter is a plane-parallel slab of transparent matter that was formerly used as an optical filter. Explain the differences of the two devices.

References

[63H1] J.J. Hopfield, D.G. Thomas, Phys. Rev. **132**, 563 (1963)
[67K1] R. Kubo, *Dynamical Processes in Solid State Optics* (Syokabo, Tokyo, 1967) and further textbooks by this author
[74G1] F. Gervais, B. Piriou, J. Phys. C **7**, 2374 (1974); Phys. Rev. B **10**, 1642 (1974); ibid. **11**, 3944 (1975)
[85C1] H.B. Callen, *Thermodynamics and an Introduction to Thermostatistics*, 2nd edn. (Wiley, New York, 1985)
[72W1] F. Wooten, *Optical Properties of Solids* (Academic, New York, 1972)
[76P1] R.W. Pohl, *Optik und Atomphysik*, 13th edn. (Springer, Berlin, 1976)
[77L1] J. Lagois, Phys. Rev. B **16**, 1699 (1977)
[81K1] R. Kuhnert, R. Helbig, K. Hümmer, Phys. Status Solidi (b) **107**, 83 (1981)
[93F1] M. Fiebig, D. Fröhlich, Ch. Pahlke-Lerch, Phys. Status Solidi (b) **177**, 187 (1993)
[01G1] B. Gil, Phys. Rev. B **64**, 201310 (R) (2001)
[01W1] M. Wegener, *Advances in Energy Transfer Processes* (1999) (World Scientific, Hongkong, 2001), p. 215
[02K1] F. Kremer, A. Schönhals (eds.), *Broadband Dielectric Spectroscopy* (Springer, Berlin, 2002)

[02S1] W. Schäfer, M. Wegener, *Semiconductor Optics and Transport Phenomena* (Springer, Berlin, 2002)
[06A1] K.C. Agarwal et al., Phys. Rev. B **73**, 045211 (2006); Phys. Status Solidi B **243**, 914 (2006)
[07B1] Y.B. Band, *Light and Matter: Electromagnetism, Optics, Spectroscopy and Lasers* (Wiley, Chichester, 2007)
[07K1] C. Klingshirn, Chem. Phys. Chem. **8**, 782 (2002); Phys. Status Solidi B **244**, 3027 (2007)
[07S1] B.E.A. Saleh, M.C. Teich, *Fundamentals of Photonics*, 2nd edn. (Wiley, Hoboken, 2007)
[10K1] C. Klingshirn et al., Phys. Status Solidi B **247**, 1424 (2010)
[10K2] C.F. Klingshirn, B.K. Mayer, A. Waag, A. Hoffman, J. Geurts, *Zinc Oxide: From Fundamental Properties Towards Novel Applications* (Springer, Heidelberg, 2010)
[10Y1] P.Y. Yu, M. Cardona, *Fundamentals of Semiconductors*, 4th edn. (Springer, Berlin, 2010)

Chapter 8
The Polariton Concept

Abstract When light propagates in matter it couples to optical excitations like optical phonons, electronic transitions, excitons etc. which leads to a polarization wave accompanying the electromagnetic wave. Quantization of this mixed propagating wave results in new quasi-particles called *polaritons*. We review here the general properties of polaritons like their dispersion relation and resulting optical properties in the Lorentz-oscillator model for (crystalline) matter including effects of anisotropy. We explicitly discuss the validity range of the polariton concept for different ratios of the distance of Lorentz oscillators and wavelength. Then we consider the consequences of coupling of the oscillators (i.e., spatial dispersion) for the dispersion, optical functions, and experimental spectra. Finally we introduce surface polaritons and their dispersion relation.

8.1 Polaritons as New Quasiparticles

In this chapter we want to elaborate in more detail what is actually propagating when "light" travels through matter. In vacuum the situation was quite clear on our present level of understanding in Sect. 2.2. Light in vacuum is a *transverse electromagnetic wave*, the quanta of which are known as *photons*. So far we have described two general approaches to describe the interaction of light with matter. For one this is the perturbative treatment or *weak-coupling* case. In this approach (see Chap. 6), the electromagnetic field and the excitations of the matter are treated as independent entities that weakly interact. Optical transitions as shown in Fig. 6.1 are described in discrete steps where a photon is absorbed (or emitted) and the matter goes from the ground state to the excited state (or the opposite way). The weak-coupling approach is sufficient for many purposes but, if we look closer, we see that this is not the whole story. The optical transition between two states of the matter is necessarily connected with some polarization density $\boldsymbol{P}$. Otherwise the transition would be optically forbidden, i.e., it would not couple to the electromagnetic field e.g. via the dipole-operator. On the other hand, we know that every oscillating polarization emits an electromagnetic wave, which may act back onto the incident electromagnetic field. This interplay is the basis of the *strong-coupling* limit between light and matter and was introduced

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_8

for a simple model system in the last chapter (Chap. 7). The propagation of a light wave in the strong-coupling regime is consequently accompanied by propagation of a polarization-density wave related to the coupled optical transition. Quantization of this mixed wave leads to the concept of *polaritons* (see e.g., [58H1, 62P1, 63H1, 75C1, 84A1, 85H1, 98K1]) which we want to elaborate in the following.

Due to the relation (see Sect. 2.3):

$$\boldsymbol{P} = \varepsilon_0 \left(\varepsilon(\omega) - 1\right) \boldsymbol{E}, \tag{8.1}$$

the electric field in matter is always accompanied by a *polarization wave*. This statement is true as long as $\varepsilon(\omega)$ or $\tilde{n}(\omega)$ deviate from one and holds in the whole spectral range from $\omega = 0$ to well beyond the highest eigenfrequencies of the material. These highest eigenfrequencies are actually situated in the X-ray region, for example, the K or L absorption edges. Above this region $\varepsilon(\omega)$ and $\tilde{n}(\omega)$ approach unity from below according to (7.19). In other words, we can state that, for the whole spectral range discussed in this book, i.e., for the IR, VIS and UV, light travelling in a solid is always a mixture or superposition of an electromagnetic wave and a "mechanical" polarization wave. This is not only true for semiconductors, but also for other dense media like metals, insulators, liquids and close to resonance even for gases. This mixed state of electromagnetic and polarization waves is quantized (see below) in the sense that it can exchange energy with other systems, including the photon-field in vacuum only by integer multiples of $\hbar\omega$.

The name of these energy quanta is *polariton* which is composed of the words polarization and photon. This name also indicates its nature of a (quasi-)particle (like phonon, plasmon etc.; an explanation of the term "quasi-" will be given in Sect. 11.1). The polarization waves of the polariton can include the motion of different lattice ions relative to each other, of two-particle (i.e., electron–hole pair) excitations, or the collective motion of the electron cloud with respect to the lattice of charged rump atoms. As will be discussed later in detail these excitations can also be quantized to form quasi-particles with energy $\hbar\omega$ and momentum $\hbar\boldsymbol{k}$ in a similar way as for photons. The names of these quanta, or quasi-particles, are *phonons*, *excitons* and *plasmons*, respectively. Even electronic excitations within the nucleus or magnetic excitations within the spin systems of electrons or nuclei (*magnons)* can form mixed states with the light field. Now it is important to repeat the above statement about what is propagating as "light" in matter, namely that it is a mixture of photons and other quasi-particles that describe the quanta of the polarization field. Please note that these mixed states are now quantized, not its constituents.

In the mechanical model used in Chap. 7 we would describe this phenomenon as follows: an incident electromagnetic wave excites the oscillators (optical excitations). This oscillation is connected with a polarization which itself radiates again an electromagnetic wave. The latter interferes with the incident one and in turn excites the oscillators, etc. This is the intuitive picture of *polariton propagation*.

The quantum-mechanical aspects of polaritons are mathematically better described in the picture of *second quantization*. We assume that the wave vector $\boldsymbol{k}$ is a good quantum number, i.e. we consider crystalline materials. The Hamiltonian reads

according to [93H1, 93P1]:

$$H = \sum_{\boldsymbol{k}} \hbar\omega_{\boldsymbol{k}} a_{\boldsymbol{k}}^{+} a_{\boldsymbol{k}} + \sum_{\boldsymbol{k}'} E(\boldsymbol{k}') B_{\boldsymbol{k}'}^{+} B_{\boldsymbol{k}'} + \mathrm{i}\hbar \sum_{\boldsymbol{k}} g_{\boldsymbol{k}} \left(B_{\boldsymbol{k}}^{+} a_{\boldsymbol{k}} + \mathrm{h.c.} \right). \tag{8.2}$$

The first two terms on the right-hand side contain the number operators (see also Sect. 3.2) $a_{\boldsymbol{k}}^{+} a_{\boldsymbol{k}}$ and $B_{k'}^{+} B_{k'}$ of the photons and of the coupled quasi-particles which are in most cases also bosons. These two terms represent the Hamiltonian of the non-interacting systems. The third term describes the interaction, e.g., the annihilation of a photon $a_{\boldsymbol{k}}$ and creation of a quasi-particle $B_{\boldsymbol{k}}^{+}$ (under momentum conservation) or vice versa. The prefactors $g_{\boldsymbol{k}}$ simply contain the transition matrix elements H_{ij} discussed in Sect. 6.2. If we would threat these terms as a perturbation, we arrive back at the weak-coupling limit.

The crucial point now is that the whole Hamiltonian (8.2) can be diagonalized by a proper choice of linear combinations $p_{\boldsymbol{k}}$ of creation and annihilation operators of photons and of the quasi-particles representing the matter excitations. In the following, we merely outline briefly this Bogoliubov-transformation-like procedure. For details the reader is again referred to the literature [93H1, 93P1].

This procedure brings the Hamiltonian of (8.2) into the following form:

$$H = \sum_{\boldsymbol{k}} E_{\boldsymbol{k}} p_{\boldsymbol{k}}^{+} p_{\boldsymbol{k}} \tag{8.3}$$

with suitable coefficients $u_{\boldsymbol{k}}$ and $v_{\boldsymbol{k}}$:

$$p_{\boldsymbol{k}} = u_{\boldsymbol{k}} B_{\boldsymbol{k}} + v_{\boldsymbol{k}} a_{\boldsymbol{k}} \,. \tag{8.4}$$

For the dependence of the magnitudes of the coefficients $u_{\boldsymbol{k}}$ and $v_{\boldsymbol{k}}$ on the energetic distance from the resonance see e.g. Sect. 21.1. The $p_{\boldsymbol{k}}$ and $p_{\boldsymbol{k}}^{+}$ are the annihilation and creation operators for the quanta of the mixed state of photon and polarization wave with wave vector $\boldsymbol{k}$. They are consequently called polaritons and $E(\boldsymbol{k})$ is their dispersion relation.

8.2 Dispersion Relation of Polaritons

All wave-like excitations can be described by two quantities, namely their (angular) frequency ω which is connected with the quantum energy simply by $E = \hbar\omega$ and their wave vector $\boldsymbol{k}$ which gives the (quasi-)momentum $\hbar\boldsymbol{k}$. The relation which connects ω and $\boldsymbol{k}$ is usually called *dispersion relation* $E(\boldsymbol{k})$ or $\omega(\boldsymbol{k})$.

The dispersion relations $E(\boldsymbol{k})$ which we have encountered so far were very simple. For photons in vacuum it was a straight line with slope $\hbar c$ (Fig. 3.1) and for the polarization wave of an ensemble of uncoupled oscillators a horizontal line (Fig. 7.2a). The relation between ω and $\boldsymbol{k}$ for polaritons, i.e., for the light quanta in matter, can be

derived from classical physics and agrees with the results of the quantum-mechanical treatment outlined above.

We recall that the wave vector in matter $\boldsymbol{k}$ is connected with the wave vector in vacuum $\boldsymbol{k}_{\mathrm{vac}}$ by the complex refractive index $\tilde{n}(\omega)$ (2.39). To get rid of the vector character, we consider the squares

$$\boldsymbol{k}^2 = k^2 = \tilde{n}^2(\omega) k_{\mathrm{vac}}^2 \,. \tag{8.5}$$

Now we also recall (2.39) and (2.16) saying that $\tilde{n}^2(\omega) = \varepsilon(\omega)$ and $\boldsymbol{k}_{\mathrm{vac}}^2 = (2\pi/\lambda_{\mathrm{vac}})^2 = (\omega/c)^2$ and obtain again (2.36)

$$\frac{c^2 k^2}{\omega^2} = \varepsilon(\omega) \,. \tag{8.6}$$

This is the so-called *polariton equation* which we used in Sect. 2.3 to describe the macroscopic properties of light propagation in matter. On the other hand, we considered a microscopic model to explicitly calculate $\varepsilon(\omega)$. The later dependence is given in the vicinity of an isolated resonance by (7.20) and (7.22). Combining (8.6) and (7.22) we find

$$\frac{c^2 k^2}{\omega^2} = \varepsilon_{\mathrm{b}} + \frac{f}{\omega_0^2 - \omega^2 + \mathrm{i}\omega\gamma} \,. \tag{8.7}$$

This is an implicit representation of $\omega(\boldsymbol{k})$ for polaritons. For the simplest case, namely vanishing damping γ and no dependence of ω_0 or f on k, it is quite easy to calculate $\boldsymbol{k}(\omega)$ and $\omega(\boldsymbol{k})$. We do not give the formulas here because they bring no further physical insight, but in Fig. 8.1 we give the dispersion relation for the case just mentioned considering only one resonance. For several resonances the r.h.s. of (7.22) has to be replaced by (7.20).

The dispersion relation starts for $\omega = 0$ and $\boldsymbol{k} = 0$ as a straight line. This part is called the *lower polariton branch* (LPB). Since the dispersion of photons in vacuum is also a straight line, but with a slope $\hbar c$ instead of $\hbar c\sqrt{\varepsilon_{\mathrm{s}}}$, the dispersion relation is said to be "photon-like" as long as it is a straight line.

The dispersion relation bends over when it approaches the resonance frequency ω_0. In this region the polariton dispersion is called e.g. phonon-like or exciton-like, depending on whether the resonance is related to a phonon or an exciton, respectively. Between the transverse and longitudinal eigenfrequencies there is no propagating mode for the present approximation of uncoupled oscillators and for vanishing damping, which is reflected in a stop band or Reststrahlen band as discussed in Sect. 7.4. There may be a longitudinal branch at $\hbar\omega_{\mathrm{L}}$, which does usually not couple to the electromagnetic field. At $\hbar\omega_{\mathrm{L}}$ also the *upper polariton branch* (UPB) begins. This one bends upwards again displaying a photon-like behavior, but now with a slope $\hbar c\sqrt{\varepsilon_{\mathrm{b}}}$ compared to $\hbar c\sqrt{\varepsilon_{\mathrm{s}}}$ for the LPB. Between ω_0 and ω_{L}, the wave-vector $\boldsymbol{k}$ is purely imaginary since $\tilde{n}(\omega)$ is imaginary in this range. The consequences for the optical properties have already been discussed in Sect. 7.4.

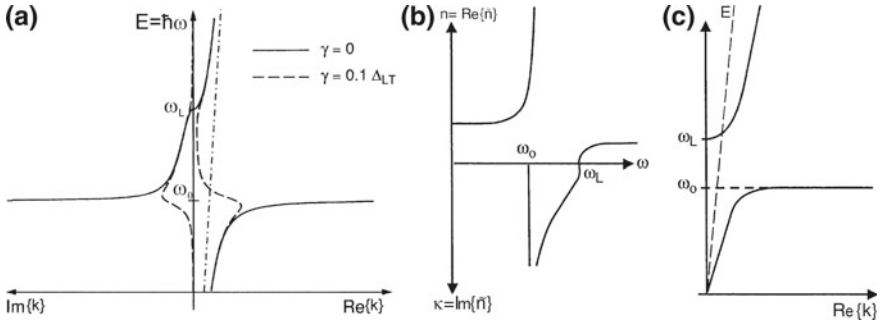


Fig. 8.1 The polariton dispersion in the vicinity of a single resonance for vanishing damping (solid lines) and finite damping (dashed lines) for $\varepsilon_b = 1$. The dashed-dotted line gives the dispersion of photons in vacuum (**a**); real and imaginary parts of $\tilde{n}(\omega)$ for vanishing damping (**b**); construction of the polariton dispersion (solid lines) from the dispersions of en ensemble of uncoupled oscillators and photons (dashed lines) by application if the non-crossing rule (here $\varepsilon_b > 1$) (**c**)

Actually, the dispersion relation shown in Fig. 8.1a can be deduced from the optical properties of the oscillator ensemble. If we take the spectral dependencies of $n(\omega)$ and of $\kappa(\omega)$, i.e., of the real and imaginary parts of the complex refractive index $\tilde{n}(\omega)$, from Fig. 7.5, turn the ω-axis from the x-direction into the y-direction (see Fig. 8.1b) and multiply $n(\omega)$ and $\kappa(\omega)$ by ωc^{-1}, i.e., essentially by a straight line through the origin, according to

$$\mathrm{Re}\{k\} = n(\omega)\omega c^{-1}; \quad \mathrm{Im}\{k\} = \kappa(\omega)\omega c^{-1}, \tag{8.8}$$

we obtain Fig. 8.1a.

The dispersion relation of the polariton can also be qualitatively understood from the quantum-mechanical *non-crossing rule*. This non-crossing rule says roughly the following: There are two energy levels E_1 and E_2, which depend on some parameter p. This parameter can be the wave vector, a particle density, a static electric or magnetic field, a strain field, etc. We assume that these two levels cross as a function of the parameter as sketched in Fig. 8.2 by the dashed lines. If there is any coupling between these two levels, then the cross-over point disappears, and the two levels repel each other in the way shown by the solid lines in Fig. 8.2. The splitting at the former cross-over point is proportional to the coupling strength between the two levels.

Actually this behavior is not only found in quantum-mechanical systems. It is basically the splitting of the eigenfrequencies of two harmonic oscillators caused by some coupling (e.g., in the case of coupled identical pendula). This example is found in every undergraduate textbook on classic mechanics including the beating if both eigenfrequencies are excited. In [98J1] it has been shown in detail that the classical and the quantum-mechanical splitting of coupled oscillators are indeed governed by the same set of equations.

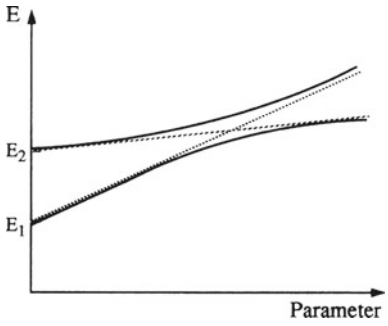


Fig. 8.2 A sketch to illustrate the quantum-mechanical non-crossing rule for two non-interacting (dashed line) and interacting (solid line) energy levels

We now apply the general understanding of avoided crossing obtained above and with Fig. 8.2 to the polariton problem (see Fig. 8.1c). The "parameter" is now obviously the $\boldsymbol{k}$ vector. The horizontal line represents the dispersion of the ensemble of uncoupled oscillators and the steep straight line the one of the photons. If the ensemble of oscillators does not couple to the electromagnetic field of the light, the dashed lines already describe the oscillator system. If there is, however, a finite coupling (e.g., a non-vanishing dipole or quadrupole coupling matrix element) the crossing is avoided and the solid lines describe the dispersion relation of the coupled system. Note that (apart from a possible longitudinal branch, not shown in Fig. 8.1c for clarity) the solid lines in Fig. 8.1c are the only states of this system. There is no state at $\hbar\omega_0$ for $\boldsymbol{k} = 0$, as we shall see later in various realistic examples.

Evidently the oscillator strength f is connected with the longitudinal-transverse splitting ((7.25a), (7.25b)) and thus a measure for the coupling strength of the electromagnetic radiation field to the excitation in the solid. Another measure is the so-called *Rabi splitting*. It is defined as the minimum vertical energetic distance between the split branches in the $E(\boldsymbol{k})$ dispersion relation in the case of coupling. Inspection of e.g. Fig. 8.1c shows immediately, that for polaritons the Rabi splitting is much larger than the longitudinal transverse splitting but both vanish for vanishing oscillator strength. We will discuss the properties of the Rabi splitting in more detail in the second volume of this book e.g. when discussing coherent phenomena or cavity polaritons.

If we relax the assumption of negligible damping, the polariton dispersion is changed as shown in Fig. 8.1a by the dashed line. Now the propagating modes on the LPB below ω_T and on the UPB above ω_L experience some damping. But, there are also propagating modes between ω_T and ω_L. To describe the propagation in such a dissipative situation we have to extend the definition of the *group velocity* v_g from (2.45) to

$$v_\mathrm{g} = \frac{\partial\omega}{\partial \mathrm{Re}\{k\}} \tag{8.9}$$

We note that this group velocity can be larger than the velocity of light $\mathrm{d}\omega/\mathrm{d}k > c$ (*superluminal propagation*) close to ω_T and it is negative $\mathrm{d}\omega/\mathrm{d}k < 0$ in a large region. In bulk material there is often no need to consider this further due to the strong

absorption which is necessarily connected with the region of negative v_g. Also for the more general case of non-vanishing spatial dispersion (Sect. 8.4) we will find that the problem of negative v_g essentially disappears. But for thin slabs of material one observes reshaping of propagating short light pulses. The group velocity is here interpreted as the propagation velocity of the pulse-envelope maximum. This maximum can leave the slab earlier than for propagation of the same distance in vacuum ($d\omega/dk > c$) or it leaves the sample earlier than the maximum of the incident pulse enters the sample ($d\omega/dk < 0$) [70G1, 82C1]. It should be noted that the fact, that both, phase- and group velocity can exceed for some frequencies the vacuum speed of light is not in conflict with special relativity and causality [70G1]. In particular, the *front velocity*, i.e. the velocity at which the first deviation from zero of the electric field of the pulse moves, can never exceed c [94R1].

We have stated several times that at $\hbar\omega_L$, i.e., at the frequency where $\varepsilon(\omega = \omega_L) = 0$, a *longitudinal branch* may exist. We want to give now some insight into when this will be the case with the help of Fig. 8.3. In Fig. 8.3a we assume that the ensemble of oscillators can be excited with the same eigenfrequency in all three independent directions of space. This situation frequently occurs to a good approximation in cubic materials. If a light beam enters the sample, e.g., in the direction k_x, then the oscillators oriented along y and z couple to the light field and form a twofold degenerate, transverse upper and lower polariton branch, while the oscillators oriented in the x direction are parallel to k_x and give rise to a non-degenerate longitudinal branch.

In case of a uniaxial material, we frequently find oscillators, which can either be excited with the same eigenfrequency in the xy plane normal to the crystallographic $c \equiv z$ axis or with another eigenfrequency parallel to z only. In the first of these two cases shown in Fig. 8.3b, namely $\boldsymbol{k}||\boldsymbol{z}||\boldsymbol{c}$, we obtain only a twofold degenerate upper and lower polariton branch, but no longitudinal branch for $\boldsymbol{k} \parallel z$ since both oscillators in the xy plane are then transversal. For $\boldsymbol{k}$ in xy plane, e.g., $\boldsymbol{k} \parallel x$ there is a nondegenerate transverse polariton branch for $\boldsymbol{E} \parallel y$ and a longitudinal branch from the x-oriented oscillators. For $\boldsymbol{k} \parallel x$ and $\boldsymbol{E} \parallel z$ the oscillators do not couple and we obtain intersecting dispersion curves.

Finally, if we consider an oscillator that can be elongated only in one direction (e.g., z) we obtain only one nondegenerate transverse lower and upper polariton branch for $\boldsymbol{k} \perp z$ and $\boldsymbol{E} \parallel z$, an intersecting longitudinal branch for $\boldsymbol{k} \parallel z$ (Fig. 8.3c) and no interaction at all for $\boldsymbol{k} \perp z$, $\boldsymbol{E} \perp z$.

If a light beam propagates with angle α different from 0° or 90° relative to the c-axis in a uniaxial material, we obtain the ordinary beam for $\boldsymbol{E}$ perpendicular to the plane defined by $\boldsymbol{k}$ and z, since $\boldsymbol{E}$ is then always in the xy plane, i.e., normal to the main section and couples to the oscillators. For $\boldsymbol{E} \perp \boldsymbol{k}$ but in the plane defined by $\boldsymbol{k}$ and z we obtain the extraordinary beam with an orientation dependent refractive index (see Sect. 5.4.1).

In the language of polaritons this is called a *mixed-mode polariton* for the following reason. The "quantization axis" for the light field is the crystallographic **c** or z axis. For a light beam propagating under an angle α to the $\boldsymbol{c}$ axis different from 0° and 90° the electric field can be decomposed into two components $\boldsymbol{E} \parallel \boldsymbol{c}$ and $\boldsymbol{E} \perp \mathbf{c}$, the relative strength of which depends on α. For the situation of Fig. 8.3b only the

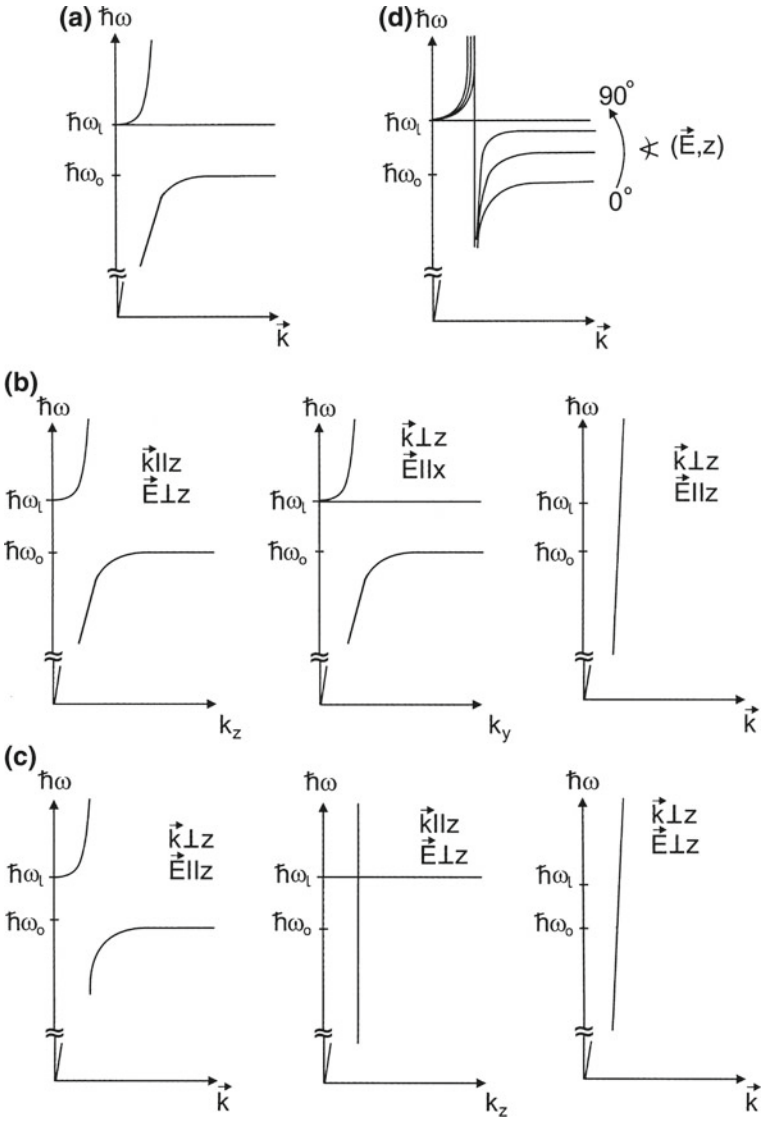


Fig. 8.3 Sketches of the polariton dispersion and of possible longitudinal branches for various orientations and an ensemble of oscillators that can be excited with the same eigenfrequency and oscillator strength in all three directions of space (**a**), only in the xy plane (**b**), and only along $z = c$ (**c**). The angle-dependent mixed-mode polariton dispersion for the situation as in **c** is shown in **d**

component $\boldsymbol{E}\perp\boldsymbol{c}$ couples to the oscillators and in Fig. 8.3c only $\boldsymbol{E}\parallel\boldsymbol{c}$. In Fig. 8.3d we show the dispersion of these mixed-mode polaritons for an ensemble of oscillators, which can be excited only in the $\boldsymbol{c}$ direction, and light polarized linearly in the $\boldsymbol{k}z$ plane. The oscillator strength and thus the $L-T$ splitting depend on the ratio $\boldsymbol{E}\parallel\boldsymbol{c}/\boldsymbol{E}\perp\mathbf{c}$. It is maximal for $\beta = \angle(\boldsymbol{E}, z) = 0°$ corresponding to the situation on the l.h.s. of Fig. 8.3c and is zero for $\angle(\boldsymbol{E}, z) = 90°$ corresponding to the middle one. In between the oscillator strength and Δ_{LT} vary according to

$$\Delta_{\mathrm{LT}}(\beta) = \Delta_{\mathrm{LT}}(\beta{=}0)\cos^2\beta\,. \tag{8.10}$$

as shown in Fig. 8.3d.

It should be noted here that we always used the angles α and β of $\boldsymbol{k}$ and $\boldsymbol{E}$ relative to $\boldsymbol{c}$ in the medium, i.e., Snell's law with a β-dependent refractive index has to be used in addition when considering the external angles.

For biaxial materials the situation becomes even more complex, but as stated in Sect. 5.4.1, this leads to crystal optics, which is beyond the scope of this book.

Similarly to the dielectric function in Fig. 7.9 the polariton dispersion is modified in the case of two close-lying resonances, i.e., the transition-matrix element cannot be deduced directly from the width of the stop band.

8.3 Validity Range of the Polariton Concept

The concept of polaritons has been described in words as early as 1946 [46B1] in nuclear paramagnetic resonance, however without using the then still unknown term "polariton". This notation was then later introduced for crystalline solids [63H1], but it can be easily generalized [84K1] to insulators, metals and non-crystalline materials including liquids or gases. Indeed, this concept is nowadays recognized not anymore as a specialized phenomenon restricted to semiconductors and to low temperatures, where the thermal broadening of the resonances is small. It is actually necessary to explain many optical phenomena in matter, even including intranuclear transitions in the X or γ-ray region.

We will illustrate the importance of the polariton concept in the following chapters by a large number of examples from semiconductor optics, e.g., for the infrared (IR) part of the spectrum (phonon-, magnon- or plasmon-polaritons) as well as for the visible (VIS) and near ultraviolet (exciton- and plasmon-polaritons). This will include in the following chapters also excitations in systems of reduced dimensionality, cavity polaritons (i.e. coupling of cavity modes to transitions in quantum emitters) and polaritons in structured materials like photonic crystals. In this section we want to draw attention to the fact that the polariton model is valid for any resonance coupling to the electromagnetic field. We will discuss some general properties of polaritons and illustrate them for various types of matter.

8.3.1 Common Optical Properties of Polaritons

To demonstrate that a spectrum of the refractive index as shown, e.g., in Fig. 7.5 is not solely characteristic of solids we present a more than 100 year-old example for Na vapor in Fig. 8.4 (for details see [76P1, 02W1, 03L1, 09M1]). If a continuous spectrum (a so-called white light spectrum) is transmitted through a cell filled with Na vapor of homogenous density (Fig. 8.4a, b) one finds in the transmitted light absorption bands. One of the most prominent bands corresponds to the well-known yellow emission (and absorption) line with is actually a double line at 589.35930 and 588.99631 nm corresponding to the 2.1 eV range. The absorption process is due to the transition of the outer electron from $^2S_{1/2}$ to $^2P_{3/2}$ and $^2P_{1/2}$, respectively [14D1]. The emission is just the reverse process.

If the white light beam is sent through a prism of Na vapor, the light is deflected for wavelengths around this resonance as shown in Fig. 8.4c, d. The behavior is just the one we expect for the refractive index and the polariton dispersion in the vicinity of a resonance. Approaching the resonance from lower photon energies, n and $\boldsymbol{k}$ increase rapidly. At the resonance absorption occurs and above the refractive index approaches $\sqrt{\varepsilon_b}$ from below. This explains the deflection of light in opposite directions for $\omega < \omega_0$ and $\omega > \omega_L$, respectively. In contrast to solids, gases and vapors are very dilute. Consequently, the oscillator strength is small and n deviates considerably from unity only in the very vicinity of the resonance. Similar experiments have been performed with many other vapors of, e.g., Sr, Ca, Ba [02W1, 03L1, 09M1]. The results allowed to develop and to verify the concept of the Lorentz oscillators. At the time of these experiments neither the concept of photons nor of polaritons was known. The quantization of the electromagnetic field in photons was first introduced by Planck in 1900 and allowed him to explain black body radiation (see (4.24)). This quantization was confirmed in the explanation of the (external) photoelectric effect by Einstein (1905). But the term "photon" and the properties of these quanta were developed only much later, e.g., by A. Compton.

From our present point of view, it is clear that the light field in Na vapor is also quantized in energy packets of $\hbar\omega$. The fact that the dispersion relation $\hbar\omega(\boldsymbol{k})$ and the refractive index deviate in the vicinity of the resonance clearly from $\hbar\omega = \hbar ck$ and $n = 1$, respectively, is proof that these quanta consist not only of electromagnetic radiation, but that they are necessarily accompanied by a polarization wave. Therefore it is straightforward to generalize the concept of polaritons beyond ordered or crystalline systems.

One point which still has to be considered, however, is the wave vector. We will come back to this aspect in Sect. 8.3.2 but state here again, that the polariton concept is now well established even in gases. It comes there under key words like "slow light", "dark polaritons" or "stationary light pulses". A small selection of references for this rapidly developing field is [00F1, 03B1, 05A1, 06G1, 06H1, 06K1, 06Z1, 10O1].

To conclude the discussion of Fig. 8.4 it should be noted that the vapor prisms were generally not obtained by keeping the Na vapor in a prism shaped cell, but that a

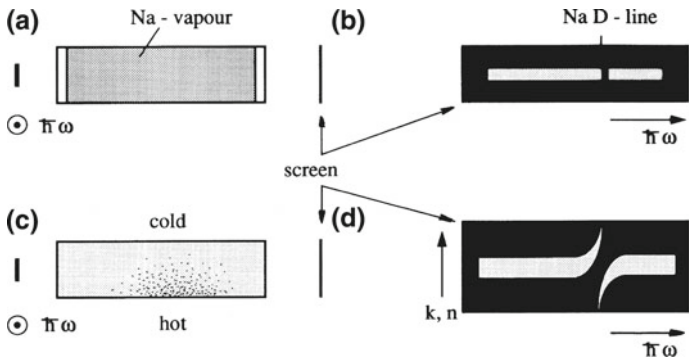


Fig. 8.4 Absorption and dispersion of Na vapor in the vicinity of the *yellow sodium line*. Schematic arrangement for the absorption (**a**) in a homogeneous vapor and the resulting absorption spectrum (**b**). Deflection in a "prism" of Na vapor (**c**) and the dispersion relation (**d**). The energy axis is normal to the paper plane in **a** and **c**; **b** and **d** are photographic positives (adapted by permission from [76P1] Springer Nature ©1976, see also [00K1])

concentration gradient has been created laterally over the light beam by a transverse temperature gradient which results in a density gradient which in turn gives a gradient in oscillator strength and finally a variation of the optical path length across the beam in a similar way as a usual (glass) prism does.

Now we present some beautiful experiments that further prove the validity of the polariton concept in seemingly completely different systems. The basic idea is illustrated in Fig. 8.5a where we show the polariton dispersion in the vicinity of a resonance. Now a short light pulse is sent onto the sample. The duration of the pulse τ_P is assumed to be so short that its spectral width $\Delta\omega$ covers the vicinity of the resonance via $\tau_P \cdot \Delta\omega \approx 1$ and $\Delta\omega > \Delta_{\mathrm{LT}}$. We can now discuss the propagation of this pulse through the medium in different ways. We know, that a pulse propagates with its group velocity $v_{\mathrm{g}} = \partial\omega/\partial k$ (see (2.45)). Evidently we find on the lower and the upper polariton branches pairs of states with equal v_{g}, i.e., equal slope as shown for one case in Fig. 8.5a. Since these states are both excited coherently by the incident pulse and since they have slightly different frequencies, they accumulate a relative phase shift, which increases linearly with time and traveling distance. The superposition of these polariton-modes pairs changes periodically from constructive to destructive and back again leading to a beating phenomenon called *propagation quantum beats*. Depending on the sample thickness and v_{g}, some of the polariton pairs will arrive at the end of the sample with constructive interference, and adjacent pairs will arrive in a destructive superposition, the next with a constructive one, etc. The consequence is that there will be no simple pulse at the end of the sample but one will find an envelope of the amplitude and intensity of the transmitted beam oscillating in time. The period of this oscillation increases with time, since the polariton pairs with large v_{g}, i.e., a steep slope of the dispersion curve arrive first. They have a bigger frequency difference and consequently a faster beating period. The pairs with smaller v_{g} also have a smaller beat period and arrive later. This is exactly the behavior

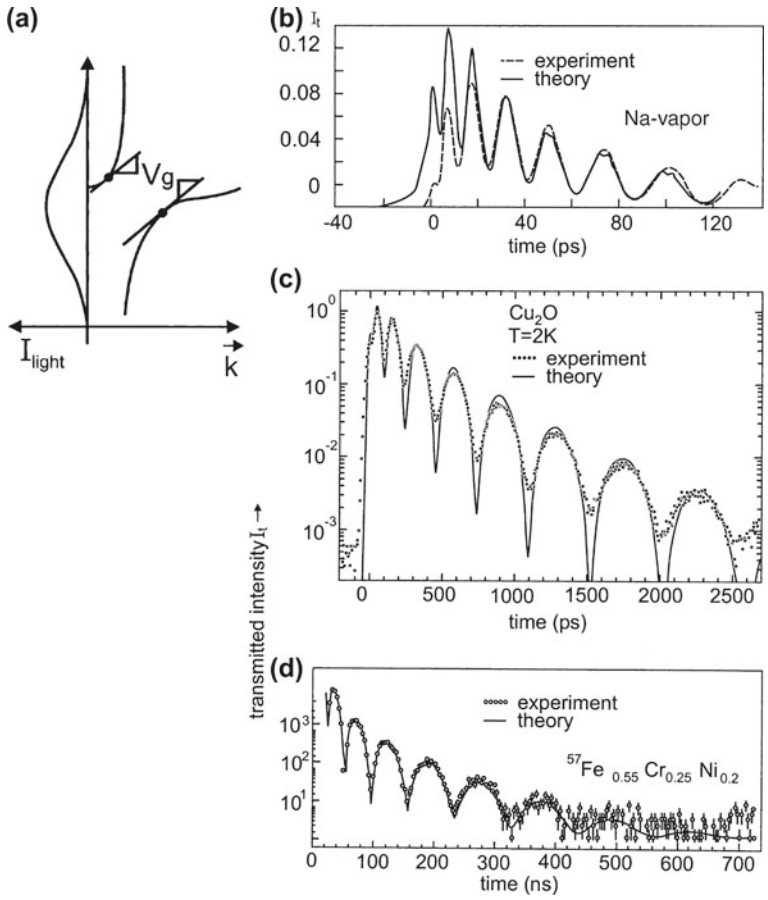


Fig. 8.5 The schematic explanation for the occurrence of *propagation quantum beats* (**a**) and their observation in Na vapor (**b**) in the ortho-exciton resonance of the yellow series in Cu_2O (**c**) and in an ensemble of ^{57}Fe nuclei (**d**) reprinted figure with permission from [84R1, 91F1] ©(1984), (1991) by the American Physical Society http://dx.doi.org/10.1103/PhysRevLett.53.552 and http://dx.doi.org/10.2298/abs1104943s, and from [99B1] Springer Nature ©(1999)

observed for an incident ns light pulse centered around the yellow resonance in Na vapor (Fig. 8.5b and [84R1]) and propagating through a cell with Na vapor with a length of 40 mm. The temporal decay of the transmitted signal is governed by the damping of the resonance and by the energy density per frequency interval of the incident pulse. The experiment in Fig. 8.5b thus confirms again the validity of the polariton concept in this system.

An alternative way to describe the experiment is to Fourier transform the incident pulse from the time to the frequency domain (see e.g. [99B1]). Then one allows each frequency component to propagate through the sample with its phase velocity according to its complex index of refraction $\tilde{n}(\omega) = n(\omega) + i\kappa(\omega)$ resulting for each

frequency ω after the sample in a well-defined phase and amplitude. All components are then coherently superimposed and this sum is transformed back to the time domain. This results in a very good description of the experimental finding as can be seen from comparison of experiment and theory in Fig. 8.5b.

Analogous propagation quantum beats are achieved also in other completely different systems as illustrated in Fig. 8.5c, d. The Fig. 8.5c shows propagation quantum beats for the exciton resonance in the yellow spectral range of Cu_2O. (More precisely, in the 1s ortho exciton of the yellow series, which is only quadrupole allowed. For details see Sect. 21.2.) The spectral region is approximately the same as for Na vapor, but Cu_2O is a crystalline semiconductor. Such type of beats have also been reported in GaAs [01N1].

The third example in Fig. 8.5d from [99B1] is, by contrast, in a completely different spectral range. The resonance is the 14.4 keV Mössbauer transition of the ^{57}Fe nucleus. The resonance is not even in the electronic system of the solid, but in the nucleus and the light pulse has been produced in a synchrotron. Nevertheless the phenomenon is identical, revealing that the polariton concept is also applicable in this spectral range of X or γ-rays.

A further common feature of the polariton dispersion is the dramatic decrease of the slope of the dispersion curve just below ω_0. Consequently, the group velocity should also obtain in this spectral region, values considerably below the vacuum speed of light $c \approx 3 \cdot 10^8$ m/s. Indeed this phenomenon has been observed, e.g., for the exciton polariton in CuCl, Cu_2O, GaAs or InSe with values of v_g down to 10^4 m/s. For more details see Sect. 21.1 or [79M1, 79U1, 81I1, 97N1, 05F1]. In the vicinity of the Mössbauer transition of ^{57}Fe nuclei at 14.4 keV, even values down to the 10 m/s range have been observed [99B1, 05S1]. Comparably low values and even a complete stop of light for a short time have been found in Na vapor [99H1, 99K1, 01K1]. In the latter case an additional trick has been used, namely excitation induced transparency in a small spectral interval of the upper of the two yellow Na lines, which results via Kramers–Kronig relations in a very steep structure in the group-velocity spectrum. Some recent references on *slow light* and polaritons in gases have already been given above.

8.3.2 *Conservation of the Photon/Polariton Momentum $\hbar k$*

In nature, there are several *conservation laws*. The most basic ones state that in a closed system, i.e., a system that has no interaction with the surroundings whatsoever, the sum of all energies E_i (including the rest-mass energies m_0c^2), of all momenta $\boldsymbol{P}_i$, of all angular momenta $\boldsymbol{L}$ (including spins) and of all electric charges ρ are constant. The sum of all entropies S_i is either constant, if there are no or only reversible processes, or it increases in the case of irreversible processes. These conservation laws are facts of experience that cannot be proven from first principles, but can be traced back to symmetries as outlined already in Sect. 5.1.3.

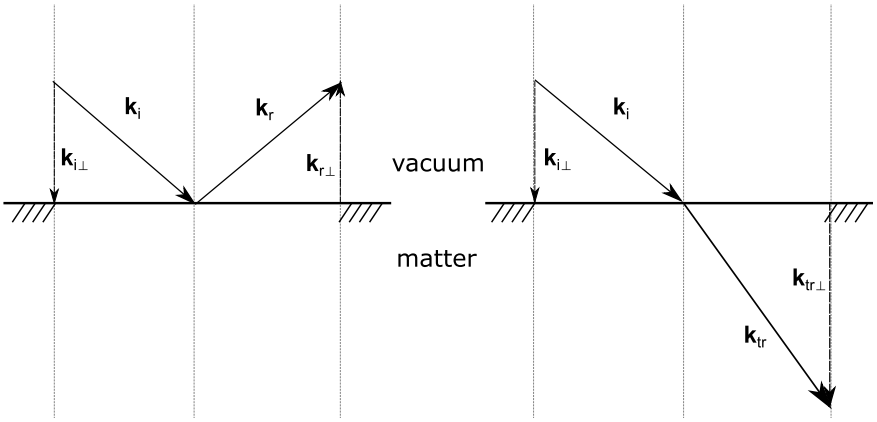


Fig. 8.6 Reflection and refraction of a single photon at an interface between vacuum and matter

The only closed system in a strict sense is the whole universe. But much smaller parts on a laboratory scale are frequently to a very good approximation closed systems. This applies, e.g., to the experiments used in lectures or lab courses to demonstrate elastic and inelastic scattering processes. We have already used a subsystem in the above sense, without mentioning it, in the derivation of the law of reflection and refraction in Sect. 5.1.2.

We want to start the discussion on momentum conservation by repeating in Fig. 8.6 the scenario of reflection or refraction of a single photon at a plane interface between vacuum and matter (compare Fig. 5.5). We argued in Sect. 5.1.3 that energy conservation requires $\omega_i = \omega_r = \omega_{tr}$ and that there is only translational invariance along the interface plane, consequently only the parallel component of $\hbar \boldsymbol{k}$ is conserved. This argument is correct for the considered subsystem. If we consider the whole closed system, the total momentum has to be conserved. This means that momentum (related to the normal components) is transferred to the matter, namely $2\hbar \boldsymbol{k}_{i\perp}$ in the case of reflection and $\hbar\,(\boldsymbol{k}_{i\perp} - \boldsymbol{k}_{tr\perp})$ for refraction. The energy transfer to the matter connected with this momentum transfer is completely negligible, however, due to the mass of the matter M. The latter is many orders of magnitude larger than the "mass" of the photon $\hbar\omega c^{-2}$.

So it is safe to say for the subsystem that the matter can "take up" momentum without energy transfer or to restrict ourselves to the subsystem "interface" with restricted translational symmetry. For a more detailed discussion of this aspect see [98B1] and the references given therein.

Now we want to use a similar set of arguments to show how conservation of $\boldsymbol{k}$ has to be treated in optics or why the law is justified that a light beam propagates along a straight line in a homogeneous medium.

We start with the situation of light scattering at a single atom in Fig. 8.7. We assume that an incident linearly polarized plane electromagnetic wave with $\boldsymbol{k}_i$ hits the atom. We consider for simplicity a single electronic transition of this atom with a

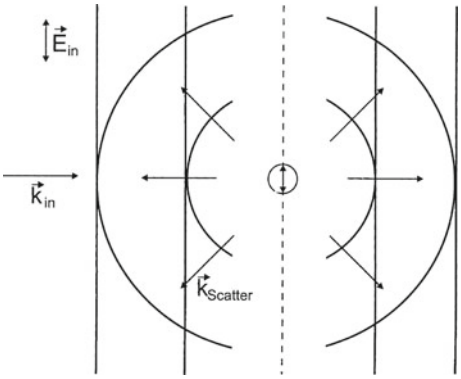


Fig. 8.7 Schematic drawing of the wave fronts in a light scattering event at a single atom or oscillator [98J1]

resonance close to the incident light frequency ω_i. In a classical approach the electron performs forced oscillations with an amplitude that depend on the detuning between ω_i and the resonance frequency ω_0 of the electronic transition and on damping. This oscillation is connected with an oscillating dipole, radiating a dipole wave as shown schematically in Fig. 8.7. Evidently, the emission goes into all directions (except the dipole axis) with the angular- dependent amplitude of an oscillating dipole. This means that a straight propagation of light interacting with a single atom is not possible.

The same result is obtained in a quantum-mechanical description. A photon comes in with $\hbar\boldsymbol{k}_i$ and $\hbar\omega_i$ and virtually excites the electron in the atom. After a time Δt, limited by $\hbar(\omega_i - \omega_0)\Delta t \leq \hbar$, where $\hbar\omega_0$ is the transition energy, the atom again radiates a photon $\hbar\boldsymbol{k}_\mathrm{s}$, $\hbar\omega_\mathrm{s}$ in an arbitrary direction. If we repeat this experiment many times, the integrated angular intensity distribution will be the one of a radiating dipole. The process is known as light scattering or if ω_i approaches ω_0 as *resonance fluorescence*.

Momentum conservation can now be discussed in the two ways outlined above. We can either state that space with a single atom, e.g., at the origin of the coordinate system, no longer has translational invariance. Consequently, there is no longer momentum conservation and an incident photon $\hbar\boldsymbol{k}_i$ can be scattered in any direction $\hbar\boldsymbol{k}_\mathrm{s}$.

Or we argue, that the law of global momentum conservation can be satisfied by the fact that the atom takes over some momentum $\boldsymbol{p}_\mathrm{atom}$

$$\boldsymbol{p}_\mathrm{atom} = \hbar(\boldsymbol{k}_i - \boldsymbol{k}_\mathrm{s})\,. \tag{8.11a}$$

The energy transfer connected with this momentum transfer

$$E_i = \boldsymbol{p}_\mathrm{atom}^2/2M \tag{8.11b}$$

is in the neV regime for the visible and therefore for most practical purposes (except for the case of many photons used for laser cooling of atoms) completely negligible. So we are again safe with both statements that either the atom takes up momentum

but no energy or that the translational invariance is broken by the presence of the atom and momentum conservation is relaxed.

This situation may change if we go to the X or γ-ray regime, where $\boldsymbol{k}$ increases. In the case of extremely narrow resonances, the frequency shift in the scattering process can then actually be measurable (see Problem 8.4).

Now we discuss the transition to a dense medium. When there are a few of such atoms in the coherence volume of the incident light, like in a dilute gas, the scattered waves interfere constructively or destructively as a function of their direction-dependent relative phases. So we obtain (momentarily) a rather complex scattering pattern, similar to a Speckle pattern, but it is averaged out over time if the scattering centers move statistically as the atoms in a dilute gas do. The *coherence volume* of light is the volume over which the light wave has a well defined phase relation. The coherence volume depends on the light source and is generally smaller for incandescent light as compared to laser light. For details see, e.g., [98B2, 07M1, 07S1].

The situation becomes simpler again if there are many oscillators or atoms in the coherence volume of light. It can be shown, that all scattered waves interfere essentially destructively except for the forward scattered waves. This means that the straight propagation of a light beam, which implies $\boldsymbol{k}$ conservation, is trivial in vacuum and develops in matter again with increasing density of oscillators N.

A lower limit for the coherence volume of light V_{coh} is given by the wavelength λ

$$V_{\text{coh}} \geq \lambda^3 \,. \tag{8.12a}$$

The condition to have many oscillators in the coherence volume of light is fulfilled in any case if the relation

$$N \cdot \lambda^3 \gg 1 \tag{8.12b}$$

is fulfilled. This is equivalent with the statement that the average distance d between the atoms or oscillators is small compared to the wavelength

$$d \ll \lambda \,. \tag{8.12c}$$

Consequently the light field cannot "resolve" the individual oscillators and a description in the sense of an *effective medium* is adequate [97S1, 97W1, 99C1]. Equation (8.12) can be also used as a definition of a *dense medium*. See also the derivation of the dipole approximation in Sect. 6.2.

To summarize we can state that in a dense medium, $\boldsymbol{k}$ conservation is recovered, including its application in the laws of reflection and refraction at the boundary of a dense medium. Still the electromagnetic wave propagating in this medium is accompanied by a polarization wave, bringing us back again to the polariton concept.

If the medium is completely ordered, i.e., a defect-free crystal, then there is only the straight propagating beam, since in such a situation we recover strictly translational invariance modulo integer multiples of the reciprocal lattice vectors, (see Sect. 9.2). If there are any inhomogeneities like defects in a crystal or inhomoge-

neous strain or fluctuations of N in a gas, there is always a small amount of light that is scattered without frequency shift, the so-called *Rayleigh scattering* .

While $\boldsymbol{k}$ conservation (modulo reciprocal lattice vector $\boldsymbol{G}$) is a strict law in perfect crystals the situation is different in disordered systems like gases, liquids, amorphous solids or alloys. For a wave with a wavelength λ that fulfills (8.12), holds the law of $\boldsymbol{k}$ conservation and we make use of it whenever we look through a glass window, onto a clear lake or through atmosphere. For shorter and shorter wavelengths, however, the conditions (8.12) are less and less fulfilled and $\boldsymbol{k}$ conversation is more and more relaxed.

This transition can be nicely followed by comparing, e.g., crystalline quartz (c-SiO_2) on one side and amorphous fused silica (a-SiO_2) or glass on the other. In the visible, (8.12) is fulfilled in both cases, light propagates along a straight line, $\boldsymbol{k}$ conservation holds and there hold well-defined laws of reflection and refraction at the surface. If we consider the X-ray regime we find $\lambda \leq d$. For c-SiO_2 $\boldsymbol{k}$ conservation modulo reciprocal lattice vectors is still valid for elastic scattering, resulting in Ewald's-construction and well-defined diffraction spots fulfilling

$$\boldsymbol{k}_i - \boldsymbol{k}_\mathrm{s} = \boldsymbol{G}\,; \quad \hbar\omega_i = \hbar\omega_\mathrm{s} \Rightarrow |\boldsymbol{k}_i| = |\boldsymbol{k}_\mathrm{s}| \ . \tag{8.13}$$

where $\boldsymbol{G}$ is a vector of the reciprocal lattice (see Sect. 9.2). In contrast, in this regime we obtain for a-SiO_2 only blurred ring structures, indicating a substantial relaxation of $\boldsymbol{k}$ conversation at these short wavelengths.

So we can state that $\boldsymbol{k}$ conservation is valid in disordered systems if (8.12) is fulfilled, but is increasingly relaxed if the wavelength gets smaller than the atomic distances.

8.4 Coupled Oscillators and Polaritons with Spatial Dispersion

For our model system we have assumed up to now zero coupling between neighboring oscillators as shown in Fig. 7.1. The consequence of this assumption is a dispersion relation which is simply a horizontal line (Fig. 7.2a). We now relax this assumption and consider a more realistic coupling between neighboring oscillators. This is shown schematically in Fig. 8.8 by adding weak springs. This concept was introduced in [58H1, 62P1, 63H1, 74A1] and elaborated in detail e.g. in [75C1, 75L1, 78B1, 78H1, 78S1, 79B1, 79S1, 81B1, 81L1, 81S1, 82O1, 82S1, 84M1, 84A1, 84H1, 84S1, 85H1, 89R1, 91R1, 03C1].

The most important consequence of this coupling is that the eigenfrequency is now a function of $\boldsymbol{k}$ as shown in Fig. 7.2b. For very long wavelength, i.e., $\boldsymbol{k} \to 0$, neighboring oscillators are still in phase and the coupling springs are not elongated. Therefore the oscillators still oscillate with the same frequency as the uncoupled ones. For decreasing wavelength, the coupling springs are elongated and increase

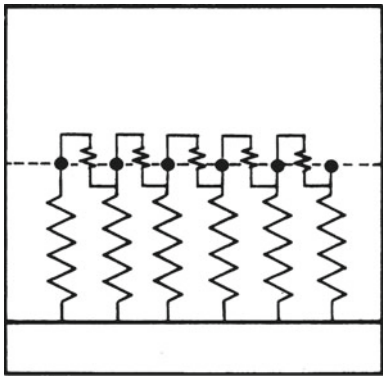


Fig. 8.8 An ensemble of coupled oscillators. Compare with Fig. 7.1 (A similar depiction was first suggested by [63H1].)

the effective spring constant. As a consequence, ω_0 increases with increasing $\boldsymbol{k}$. The resulting band width 2B indicated in Fig. 7.2b is directly proportional to the coupling strength between neighboring oscillators.

A wave packet created by elongating one or a few oscillators, as in Fig. 7.1d, now propagates with a finite group velocity $v_g = \mathrm{d}\omega/\mathrm{d}k$ and shows some dispersion. This means that the width of the spatial envelope of the wave packet increases with time as indicated in Fig. 2.1b. This phenomenon can be easily observed by throwing a stone onto the still surface of a lake. The expanding ring-like wave structure shows a significant increase of the width of its envelope function during propagation. The detailed shape of $\omega(\boldsymbol{k})$ depends on the physical nature of the oscillators and the coupling mechanism.

For our model system, ω_0 increases with $\boldsymbol{k}$ and we shall see that this is also true, e.g., for excitons. For optical phonons one has usually a decrease of the eigenfrequency with increasing $\boldsymbol{k}$. These points will be presented in more detail in following chapters. In [72B1] this problem has been treated in terms of a response function of optical phonons in Raman scattering. For the following discussion of our model system we shall use the dispersion relation of Fig. 7.2b, but the conclusions will be qualitatively valid also for other dependences of $\omega_0(\boldsymbol{k})$.

The fact that the eigenfrequency ω_0 of some excitation of a solid depends on $\boldsymbol{k}$ is often called *spatial dispersion*, for reasons given later. However, the $\omega(\boldsymbol{k})$ dependence of the photons themselves, or of the polaritons, is not usually classed as spatial dispersion.

Since we have used the word *dispersion* now in various connections, we shall summarize the meanings here:

The term "dispersion relation" or simply "dispersion" is used for the relation $E(\boldsymbol{k})$ or $\omega(\boldsymbol{k})$ for all wave-like excitations independent of the functional dependence. It can be simply a horizontal line, a linear or a parabolic relation, or something more complicated. Every excitation which has a wave-like character has a dispersion relation.

The term "spatial dispersion" means that the eigenfrequency of one of the elementary excitations in a solid depends on $\boldsymbol{k}$ and is not just a horizontal line (Fig. 7.2a, b).

In technical optics, dispersion often refers more specifically to the dependence of the refractive index n on the wavelength. Materials such as the glass used to make lenses, generally have a decreasing refractive index with increasing wavelength i.e., $\mathrm{d}n/\mathrm{d}\lambda < 0$. This behavior is called "normal dispersion" and is the usual behavior in the transparent spectral region, while an "anomalous dispersion", i.e., $\mathrm{d}n/\mathrm{d}\lambda > 0$ is limited to frequencies close to a resonance (Fig. 8.1).

Finally, the word dispersion is also used for the fact that the envelope of a wavepacket, e.g., of a short light pulse in matter, becomes spatially broader with time.

The last two meanings of the word dispersion are actually consequences or special examples of the general definition of the term given first. If not stated otherwise, we use the term "dispersion (relation)" in this book for relations $E(\boldsymbol{k})$ or $\omega(\boldsymbol{k})$.

8.4.1 Dielectric Function and Dispersion Relation of Polaritons with Spatial Dispersion

The dielectric function given in (7.22) for the simple case of an isolated resonance has to be modified if we want to take spatial dispersion effects into account. The eigenfrequency ω_0 has to be replaced by $\omega_0(\boldsymbol{k})$ and the oscillator strength f and the damping γ may also depend on $\boldsymbol{k}$ resulting in

$$\varepsilon(\omega, \boldsymbol{k}) = \varepsilon_\mathrm{b} + \frac{f(\boldsymbol{k})}{\omega_0^2(\boldsymbol{k}) - \omega^2 - \mathrm{i}\omega\gamma(\boldsymbol{k})} \,. \tag{8.14}$$

The most significant change is the fact that ε is now a function of two independent variables, ω and $\boldsymbol{k}$.

Along with the transverse eigenfrequency $\omega_0(\boldsymbol{k})$ the longitudinal eigenfrequency ω_L defined as $\varepsilon(\omega_\mathrm{L}) = 0$ also becomes $\boldsymbol{k}$-dependent, i.e., $\omega_\mathrm{L} = \omega_\mathrm{L}(\boldsymbol{k})$. The same is true for the longitudinal-transverse splitting Δ_LT which is connected with $f(\boldsymbol{k})$. In principle all of the above quantities have to be given as functions of $\boldsymbol{k}$.

Usually this situation is described—if possible—by an analytic expression for $\omega_0(\boldsymbol{k})$ and by still considering f and γ as $\boldsymbol{k}$-independent. There is clear experimental evidence that f and γ depend on $\boldsymbol{k}$ even for dipole-allowed transitions (see e.g. [82O1, 84M1, 84K1, 84S1, 89R1, 91R1]). Still these dependences are usually less critical for the correct description of the optical properties of semiconductors than the $\boldsymbol{k}$-dependence of ω_0 and will be neglected in the following.[1]

For $\omega_0(\boldsymbol{k})$ we again use for simplicity a parabolic relation, which resembles nicely the onset of the dispersion relation in Fig. 7.2b but also that of excitons (see Chap. 20).

[1] For electric quadrupole transitions, the matrix element varies linearly with $\boldsymbol{k}$. Compare (6.23), which has been truncated after the constant term for the dipole approximation.

But we stress once more that the consequences are qualitatively similar for other relations. The dielectric function now reads

$$\varepsilon(\omega, \boldsymbol{k}) = \varepsilon_\mathrm{b} + \frac{f}{\omega_0^2 + 2\omega_0 A\boldsymbol{k}^2 - \omega^2 - \mathrm{i}\omega\gamma} \tag{8.15a}$$

with

$$\omega_0^2(\boldsymbol{k}) = \left(\omega_0 + A\boldsymbol{k}^2\right)^2 = \omega_0^2 + 2\omega_0 A\boldsymbol{k}^2 + A^2\boldsymbol{k}^4 \approx \omega_0^2 + 2\omega_0 A\boldsymbol{k}^2$$
$$\text{for} \quad |\boldsymbol{k}| \ll \pi/a \ . \tag{8.15b}$$

This approximation is usually valid for massive and for effective-mass particles (see Sect. 14.5) like excitons.

To determine the dispersion relation of the polariton we combine the polariton equation with the dielectric function resulting again in an implicit relation for $\omega_0(\boldsymbol{k})$:

$$\frac{c^2\boldsymbol{k}^2}{\omega^2} = \varepsilon_\mathrm{b} + \frac{f}{\omega_0^2 + 2\omega_0 A\boldsymbol{k}^2 - \omega^2 - \mathrm{i}\omega\gamma} \ . \tag{8.15c}$$

We consider first the case of vanishing damping in Fig. 8.9a and begin with the real part of $\boldsymbol{k}$. The transverse lower polariton branch (LPB) starts photon-like and then bends over to asymptotically approach the parabolic dispersion relation of the resonance. If there is a constant Δ_LT, i.e., $f \neq f(\boldsymbol{k})$, then the longitudinal branch shown by the dashed line in Fig. 8.9a starts at ω_L and is then essentially parallel to the exciton-like part of the LPB. At ω_L the transverse upper polariton branch (UPB) also begins, going over into a photon-like (i.e., linear) dispersion relation with slightly steeper slope, as shown already in Fig. 8.1a. The imaginary part of $\boldsymbol{k}$ starts at ω_L and bends downwards to reach asymptotically a curve which is produced by reflecting the transverse eigenfrequency $\omega_0(\boldsymbol{k})$ through the point $(\omega_0, \boldsymbol{k} = 0)$. This means that the UPB has a purely imaginary continuation below ω_L.

If we include a small but finite damping the situation shown in Fig. 8.9b arises. Compare to Fig. 8.1a for the situation without spatial dispersion. It can be seen that the LPB and the UPB now extend over the whole energy range. They have both a real and an imaginary part at all frequencies. The imaginary part of the LPB is very small below ω_0, peaks between ω_0 and ω_L, and decreases gradually for higher energies. The UPB has a very small imaginary part above ω_L which increases continuously below, but there is also a small real part of the UPB below ω_L. The problem of negative group velocity in the spectral region of the resonance, mentioned already in Sect. 8.2, has become less severe and appears only in the form of a small hump in the dispersion relation for finite damping.

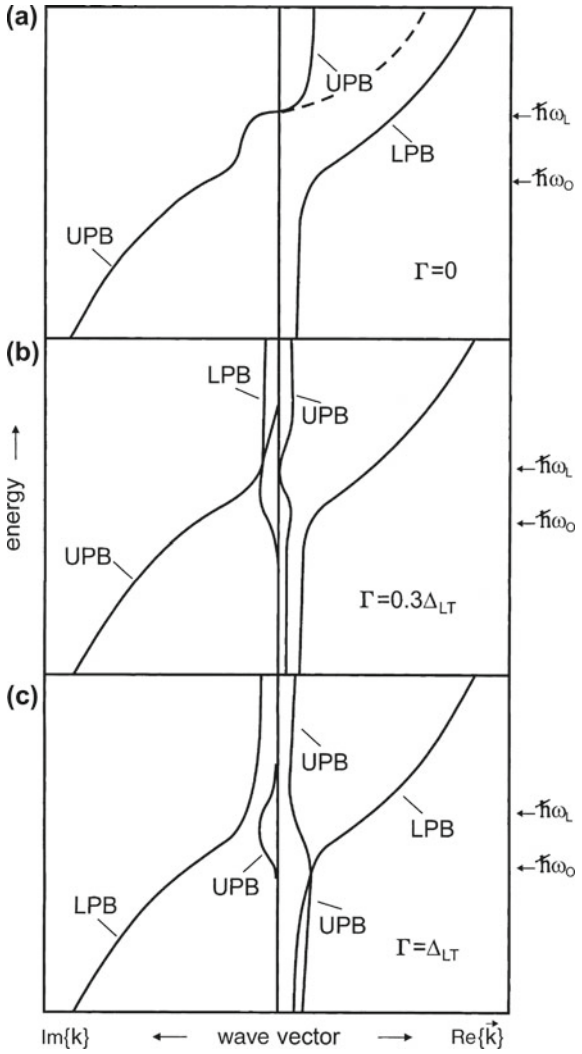


Fig. 8.9 The real and imaginary parts of a polariton dispersion relation in the vicinity of a resonance with spatial dispersion for vanishing (**a**) small $\hbar\Gamma < \Delta_{LT}$) (**b**) and strong $\hbar\Gamma \approx \Delta_{LT}$) damping (**c**) (adapted figure with permission from [84M1] ©1984 by American Physical Society https://doi.org/10.1103/PhysRevB.29.3362)

For the case of strong damping shown in Fig. 8.9c the role of the LPB is gradually taken over by the UPB below ω_0 [84M1] and one approaches a situation similar to the case without spatial dispersion (compare with Fig. 8.1a). In other words this means that one can describe the optical properties of a strongly damped resonance by neglecting the influence of spatial dispersion. However, it should be stressed, that this effect is still present in principle, but the strong damping is what dominates the optical properties.

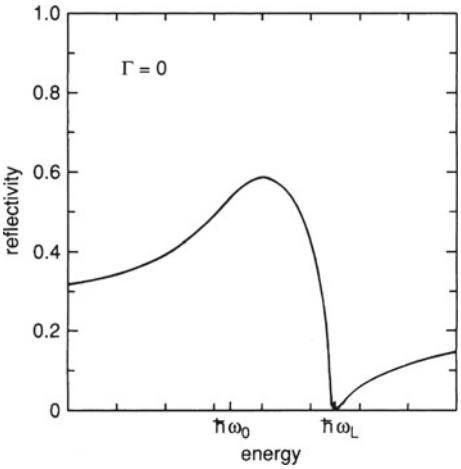


Fig. 8.10 Sketch of the reflection spectrum in the vicinity of a resonance with spatial dispersion and without damping. Compare with Fig. 7.5 (after [82R1])

Even with strong damping there remains at least one transverse branch that extends over the whole Brillouin zone in contrast to Fig. 8.1a for $\gamma \neq 0$ and possibly also a longitudinal one. These branches are the ones observed, e.g., in inelastic neutron scattering when determining the dispersion relation of TO- and LO-phonons over the whole Brillouin zone.

8.4.2 Optical Spectra and Additional Boundary Conditions

From the polariton dispersion with spatial dispersion shown in Fig. 8.9 we can easily recover the real and imaginary parts of $\tilde{n}$ by just reversing the procedure given by (8.8). We will not spend time on this procedure which is described in e.g. [75L1, 78B1, 78H1, 78S1, 79B1, 81L1, 82S1]) but discuss directly the optical spectra. In particular we will elaborate the consequences of spatial dispersion on the reflection as compared to the situation of Fig. 7.5 where spatial dispersion was still neglected. The two most important points are, first, that there is no more strict stop band, i.e., there is at least one propagating mode for every frequency ω (with and without damping). Secondly, in some spectral regions there is more than one propagating mode. This second point is especially obvious for $\omega > \omega_L$ in Fig. 8.9a.

We now want to discuss the consequences of these two new phenomena arising from spatial dispersion on the optical spectra, especially on the reflection spectrum. The fact that there is at least one propagating mode—generally with real and imaginary part—for all frequencies even between ω_T and ω_L, means that the reflectivity for (normal) incidence no longer reaches unity, even for the case of vanishing damping, as sketched in Fig. 8.10.

The fact that there is more than one mode (propagating or evanescent, i.e., $\mathrm{Im}(\boldsymbol{k}) \ll \mathrm{Re}(\boldsymbol{k})$ or $\mathrm{Im}(\boldsymbol{k}) \gg \mathrm{Re}(\boldsymbol{k})$, respectively) in the solid for one frequency

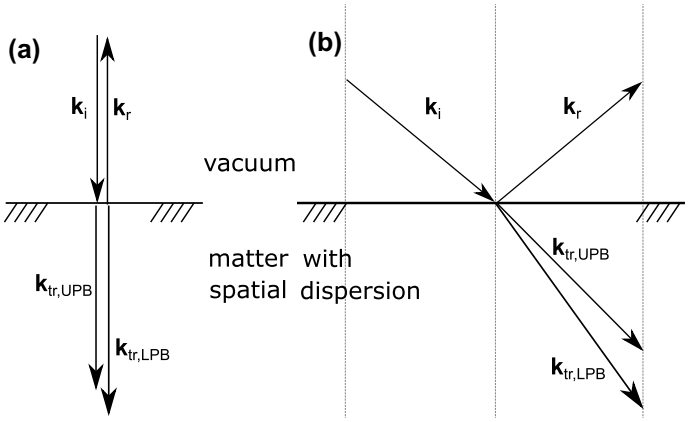


Fig. 8.11 The incident and transmitted wave vectors for $\omega > \omega_L$ of a resonance with spatial dispersion. Note that there is more than one propagating wave in the medium for both normal **(a)** and oblique incidence **(b)**

means that the two independent boundary conditions deduced from Maxwell's equations in connection with (5.6), (5.7) and Figs. 5.1, 5.2 are no longer sufficient. For a given incident wave we can deduce the amplitudes of one reflected and one transmitted wave. If there are two or even more states in the medium at the same frequency coupling to the incident field, we need one or more *additional boundary conditions* (abc).

To make the situation clear, we show in Fig. 8.11 the wave vectors for such a case for normal and oblique incidence. The incident and reflected beams obey the usual law of reflection, their components parallel to the surface are equal. The same is true for the transmitted beams, but the total length of the wave vectors on the LPB and the UPB are different, in agreement with Fig. 8.9. Obviously the two beams are travelling in different spatial directions and this is a reason why the $\boldsymbol{k}$-dependence of ω_0 is called "spatial dispersion".

Although the situation shown in Fig. 8.11 has some similarity with the one of birefringence (see Sect. 5.4.1) we point out that the physical explanation is quite different. In Fig. 5.12 the o and eo beams have orthogonal polarization and the phenomenon requires uniaxial or lower symmetry. Spatial dispersion occurs independent of the symmetry and even for cubic symmetry. Further, the two beams are polarized in the same direction. Spatial dispersion also occurs for crystals of lower symmetry which then may show birefringence in addition. In this case the dispersion curves of Fig. 8.9 have to be plotted twice with different parameters for the o and the eo beams.

As already mentioned, the abc cannot be deduced from Maxwell's equations. Their capacity is exhausted with one reflected and one transmitted beam. The complex index of refraction around the resonance is rather different for the LPB and the UPB, which therefore contribute differently to the reflection spectrum according to (5.19). As a consequence the abc should contain information about the "branching ratio",

i.e., which fractions of the incident beam couple in the medium to the LPB and to the UPB as a function of frequency.

The abc are somewhat arbitrary (we shall explain later on why) and are based mainly on arguments of physical plausibility. On the vacuum side of the interface the polarization density is trivially zero. To avoid an non-physical discontinuity in the polarization density, one possible abc is that the polarization density of the medium must be zero at the interface and develops a finite value in the medium only away from the interface z

$$\boldsymbol{P}(z=0)=0\ . \tag{8.16a}$$

Another argument says that the polarization should vary smoothly across the interface, implying that the derivative with respect to the normal direction has to vanish, resulting in

$$\left.\frac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}z}\right|_{z=0}=0\ . \tag{8.16b}$$

Some authors favor a linear combination of the two conditions:

$$\boldsymbol{P}|_{z=0}+\beta\left.\frac{\partial\boldsymbol{P}}{\partial z}\right|_{z=0}=0\quad\text{with}\quad -1\le\beta\le+1\ . \tag{8.16c}$$

The reflection spectrum shown in Fig. 8.10 is actually calculated for an exciton resonance using the material parameters of CdS and the abc (8.16a), the so-called *Pekar–Hopfield abc*. For further details of the abc and for some controversial discussions see also Sect. 21.1 and the literature on exciton polaritons given in the bibliography of this chapter.

It turns out that experimentally observed spectra, e.g., of exciton resonances, can be fitted with all the above-mentioned abc, but with slightly different values for the other parameters, such as f and γ, that describe the resonance. As a rule of thumb, one can state for all abc and weak damping that the light propagating in matter at frequencies ω sufficiently below the transverse eigenfrequency and above the longitudinal one travels almost completely on the LPB and on the UPB, respectively. "Sufficiently" means in this context $\hbar\left|\omega-\omega_{\mathrm{T,L}}\right|\gtrsim 10\Delta_{\mathrm{LT}}$. The crucial spectral range where spatial dispersion and the problem of abc are of importance is thus the resonance and its vicinity.

Now let us have a look at the transmission, including spatial dispersion. Since both the LPB and the UPB have substantial imaginary parts in the vicinity of the resonance, we again expect a dip in the transmission in the region around the resonance. At considerably lower frequencies, the light couples almost completely to the LPB—which has an almost negligible imaginary part in this region—so that the sample is transparent. The same is true significantly above ω_{L} for the UPB, at least if there are no other resonances. The shape of the effective absorption coefficient $\alpha(\omega)$ or of $\kappa(\omega)$ can look somewhat different from Fig. 7.5 possibly developing a spike at the longitudinal eigenfrequency, as indicated in Fig. 8.12.

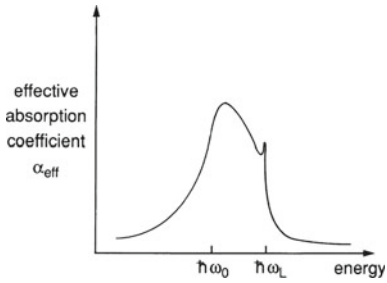


Fig. 8.12 A possible spectrum of the effective absorption coefficient in the vicinity of a resonance with spatial dispersion [82R1]

In addition, one can see from Fig. 8.9 that the imaginary parts of $\boldsymbol{k}$ and thus κ or α are different for the two polariton branches. The amplitudes and light intensities of each polariton branch therefore decay exponentially with propagation length, but the sum of both, which is the only experimentally accessible quantity, may show a non-exponential spatial decay. Furthermore the damping may in some cases be higher close to the surface compared to the bulk of the sample due to imperfections introduced into the lattice close to the surface. Strong damping, i.e., $\hbar\gamma \geq \Delta_{\mathrm{LT}}$, reduces the importance of the influence of spatial dispersion on the optical spectra as already mentioned above. The resonance in the reflection spectrum is then already so strongly washed out by damping that the details, whose description relies on spatial dispersion, are no longer observable.

It is physically not completely satisfying that the choice of the abc is to some extent arbitrary (8.16a)–(8.16c). The parameters of the resonance deduced with the use of different abc are (slightly) different. On the other hand, the value of quantities such as f cannot depend on the abc chosen by the physicist running the fitting program. The answer is that the problem of the abc is an artificial one. The dielectric function $\varepsilon(\omega, \boldsymbol{k})$ describes the optical properties in the bulk of the sample. In our derivation of $\varepsilon(\omega, \boldsymbol{k})$ in Chap. 7 no surface was included. Later on, however, we use this dielectric function to describe the optical properties of the interface between two semi-infinite half-spaces, one usually being vacuum, the other the semiconductor under consideration. The price we have to pay for this improper use of the dielectric function is the problem of the abc.

If we were to calculate the optical properties of a half-space from the beginning, the problem of the abc could be avoided. Indeed, some calculations have used this idea [79S1, 81B1, 81S1, 82S1] or [98H1, 99H2]. However, this procedure has its own problems. For example, the band-structure and the exciton states have to be calculated for the half-space, which means that we can in principle no longer use Bloch's theorem for the direction normal to the interface. This causes complications which can be overcome only with difficulty and by using various approximations and simplifications. Therefore most authors prefer to use the bulk dielectric function and some of the abc to evaluate the optical spectra. A very interesting solution to the problem of abc's was put forward by [98H1], which triggered some discussion [99H2], but allows one to avoid the abc-problem. The relation between microscopic

theory and abc has been treated in [01S1, 02M1, 04S1]. More information about the problem of abc and the rather lengthy formulas for calculating the spectra of reflection can be found in the literature given below.

8.5 Real and Imaginary Parts of Wave Vector and Frequency

Until now, we have assumed that we can describe the light wave or polariton wave propagating in matter by a real frequency and a wave vector which has a real and an imaginary part according to the complex index of refraction. In principle, however, one could take $\boldsymbol{k}$ to be real and introduce a complex frequency $\tilde{\omega}$ by including the damping term $\mathrm{i}\omega\gamma$ as in (7.14) and (7.22):

$$\tilde{\omega} = \omega - \mathrm{i}\gamma \tag{8.17}$$

At first glance, there is no reason to prefer one approach over the other. Actually both approaches are possible and it is the experiments performed that decide which model is the more appropriate one.

If we shine a monochromatic wave with well-defined frequency ω (e.g., a spectrally narrow laser beam) on the sample, or if we select such a frequency with a monochromator, then we have the situation of (7.14) and (2.40) i.e., a forced oscillation, and here we have to use a purely real ω, but a complex $\boldsymbol{k}$ to describe the decrease in amplitude of the polariton wave as it travels through the crystal.

If, on the other hand, we could by some means create at a certain time (e.g., $t = 0$), a polariton wave in the sample with constant amplitude everywhere and let it evolve for $t > 0$, then the amplitude would decay with time, but would remain the same everywhere in space. This is just the situation described by an only real $\boldsymbol{k}$ and a complex $\tilde{\omega}$. The reader might think that a wave with constant amplitude from $-\infty$ to $+\infty$ in space is highly unphysical, but in fact this approximation is as valid as that of a strictly monochromatic wave, which necessarily endures from $t = -\infty$ to $t = +\infty$. Everything with a finite temporal duration has a finite spectral halfwidth.

These are simply consequences of the fact that the time t and frequency ω domains and the space $\boldsymbol{r}$ and wave vector $\boldsymbol{k}$ domains are connected with each other by one and three-dimensional Fourier transformations, respectively. Multiplication of ω or $\boldsymbol{k}$ with $\hbar$ then immediately gives the *uncertainty relations* for energy and time or momentum and space (6.2).

Since the experimental situation discussed first is much more common than the second one, we will restrict ourselves for the rest of this book to the situation of complex $\boldsymbol{k}$ and real ω. We bear in mind, however, that in principle one needs to use both a complex $\boldsymbol{k}$ and a complex ω since all excitations usually have a finite lifetime or phase-coherence time and this is what we used in the derivation of e.g. the dielectric function in (7.6)–(7.14). Dynamic effects will be considered in more detail in the second volume of this book.

8.6 Surface Polaritons

For almost every wave-like excitation in the bulk of a solid or even liquid sample, there exists a surface or interface mode including Rayleigh-waves in earthquakes or the usual waves on the surface of water. Important applications rely on, e.g., surface acoustic phonons, surface plasmons, etc. (see, e.g., [91K1] or [88R1, 98B2, 07S2, 10S1, 15O1]). Here we want to focus on *surface polaritons* already mentioned briefly in Sect. 2.3.

Surface polaritons are also quanta or quasiparticles of the mixed state of an electromagnetic and a polarization wave. They are distinguished from bulk modes by the fact that they can only propagate along the interface between two different media. The amplitudes decay exponentially with distance from the interface on both sides, as shown schematically in Fig. 8.13a. This, surface polaritons are evanescent waves on both sides of the interface, in contrast to the one-sided evanescent wave in the case of total internal reflection of Fig. 5.3a. See also Fig. 7.4. For every volume polariton there exists a surface polariton.

We want to discuss here the conditions for which surface polaritons can exist in a general way, i.e., without specifics of the corresponding bulk polaritons. Examples for phonon-, plasmon- or exciton surface polaritons will be given in the respective chapters. For simplicity we restrict our discussion to the case of vanishing damping and no spatial dispersion: $\gamma = 0$ and $A = 0$ in (8.15). We assume that the interface is formed by an essentially non-dispersive medium I described by a constant real index of refraction $n_\mathrm{I}^2 = \varepsilon_\mathrm{I}$ on one side (for vacuum $n_\mathrm{I} = 1$) and the medium under consideration with $\varepsilon(\omega) = \tilde{n}^2(\omega)$ on the other (medium II). If the surface polariton cannot propagate into medium I or II, there must be some physical reasons preventing its decay by radiating into the half-spaces I and II.

As a first condition we may state that in medium II there are no propagating waves between the transverse and longitudinal eigenfrequencies ω_T and ω_L, as discussed for example in connection with Figs. 7.3, 7.5 and 8.1a. The propagation into medium I can be excluded if the wave vector $\boldsymbol{k}_\mathrm{s}$ of the surface polariton, which is directed

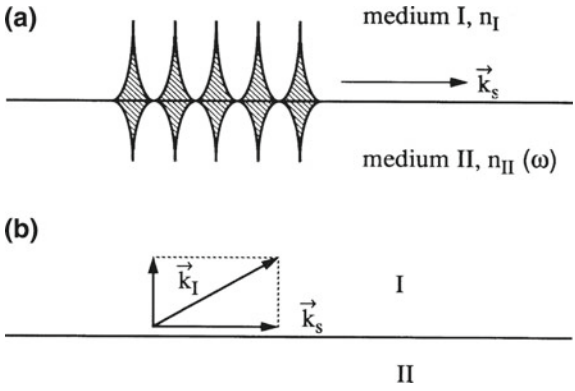


Fig. 8.13 Schematic drawing of the decay of the amplitudes of a surface polariton propagating along an interface (**a**) and a sketch to illustrate the derivation of (8.18) (**b**)

parallel to the interface, is larger than $\boldsymbol{k}_\mathrm{I}$ of any wave propagating in medium I. Under such a condition the conservation law for $\boldsymbol{k}$ parallel to the interface results with Fig. 8.13b in

$$\boldsymbol{k}_\mathrm{s}^2 + \boldsymbol{k}_\perp^2 = \boldsymbol{k}_\mathrm{I}^2 \quad \text{and} \quad \boldsymbol{k}_\mathrm{s}^2 > \boldsymbol{k}_\mathrm{I}^2 \Rightarrow \boldsymbol{k}_\perp^2 < 0\,. \tag{8.18}$$

The r.h.s of (8.18) simply says that $\boldsymbol{k}_\perp$ is purely imaginary and this is the condition for an evanescent wave normal to the interface.

We can summarize these arguments by saying that surface polaritons with frequency ω_s and wave vector $\boldsymbol{k}_\mathrm{s}$, can be expected in the spectral region given by

$$\omega_0 \equiv \omega_\mathrm{T} \leq \omega_\mathrm{s} \leq \omega_\mathrm{L} \quad \text{in medium II and} \tag{8.19a}$$

$$\boldsymbol{k}_\mathrm{s} \geq n_\mathrm{I}\omega c^{-1} \quad \text{in medium I} \tag{8.19b}$$

or

$$\mathrm{Re}\{\varepsilon_\mathrm{II}(\omega)\} < 0 \quad \text{and} \quad |\mathrm{Re}\{\varepsilon_\mathrm{II}(\omega)\}| > \varepsilon_\mathrm{I}\,. \tag{8.19c}$$

The dispersion relation of surface polaritons $\omega_\mathrm{s}(\boldsymbol{k}_\mathrm{s})$ can be deduced for the present assumptions from the boundary conditions given by Maxwell's equations. We will not go through the procedure here, but merely give the result and refer the reader for its derivation to the literature e.g. [74O1, 81L1, 82M1].

$$\boldsymbol{k}_\mathrm{s} = \left(\frac{\varepsilon_\mathrm{I} \cdot \varepsilon_\mathrm{II}(\omega)}{\varepsilon_\mathrm{I} + \varepsilon_\mathrm{II}(\omega)}\right)^{1/2} \frac{\omega}{c}. \tag{8.20}$$

We should note that $\varepsilon_\mathrm{II}(\omega)$ is negative in the region of (8.19a). In order to get a real value of $\boldsymbol{k}_\mathrm{s}$ evidently, in addition to (8.19a) and (8.19b), also (8.19c) has to be fulfilled.

The polarization of the surface polaritons is as follows. If the interface is the xy-plane and the surface wave propagates in the x-direction, i.e., $\boldsymbol{k}_\mathrm{s} \| x$ then the electric field is in the xz-plane and the magnetic induction $\boldsymbol{B}$ is oriented along the y-axis.

For large $\boldsymbol{k}_\mathrm{s}$ one finds that the surface polariton occurs at a frequency ω_s where

$$\mathrm{Re}\,\{\varepsilon_\mathrm{II}\,(\omega_\mathrm{s})\} = -1 \quad \text{for} \quad k_\mathrm{s} \gg \frac{\omega}{c} \quad \text{and} \quad \varepsilon_\mathrm{I} = 1\,. \tag{8.21}$$

In the case of a simple plasmon resonance the dielectric function reads

$$\varepsilon(\omega) = \varepsilon_\mathrm{b} + \frac{\omega_\mathrm{PL}^2}{-\omega^2 - i\omega\gamma} \tag{8.22}$$

as will be shown later in Sect. 19.1. For simplifying conditions, $\gamma \Rightarrow 0$ and $\varepsilon_\mathrm{b} = 1$ (the latter being realistic only for some simple metals), we find from (8.22) the

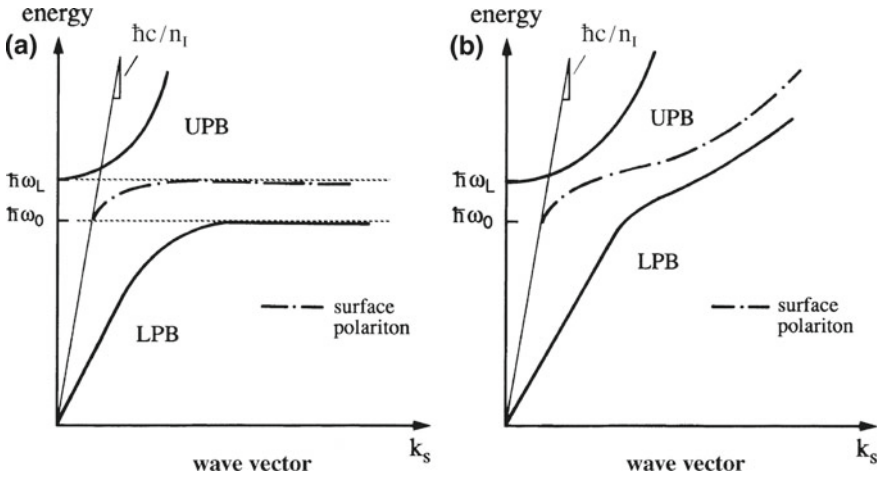


Fig. 8.14 The dispersion of the surface polariton for a resonance without (**a**), and with (**b**), spatial dispersion. The damping is assumed to be negligible in both cases

frequently given relation [80R1, 88R1]:

$$\omega_s = \omega_{PL}/\sqrt{2} \quad \text{for large} \quad \boldsymbol{k}_s \,. \tag{8.23}$$

In Fig. 8.14 we show as a summary of this chapter the dispersion of polaritons and of surface polariton for a resonance without (a), and with (b), spatial dispersion and vanishing damping. In the latter case the calculated dispersion relation $\omega_s(\boldsymbol{k}_s)$ is slightly influenced by the abc used. The condition (8.21) is reached for large $\boldsymbol{k}_s$ generally slightly below the longitudinal branch as seen, e.g., in Fig. 7.3. The dispersion relation for large $\boldsymbol{k}_s$ is parallel to the longitudinal branch without reaching it. For more details see [74O1, 81L1].

Experimental access to surface polaritons in given, e.g., by the method of attenuated total reflectance (ATR). We will explain this method and given examples of experimental data in Sect. 21.1.5.

Problems

8.1. Which dispersion relations would you expect for the polariton resulting from oscillators with the dispersion relations of Fig. 7.2a, b? Do not forget that a finite coupling between photons and the oscillator necessarily implies a finite Δ_{LT}.

Check if you were right when you come to the chapters on phonon and on exciton polaritons.

8.2. Calculate the frequency shift a photon experiences when it is scattered off an atom in a backward direction. Compare with the homogenous linewidth of luminescence lines in semiconductor optics, which hardly fall below 0.1 meV.

8.3. Inform yourself about the possibility of cooling atoms by absorption and emission of photons.

8.4. What is the Mösbauer effect? How does it work?

8.5. In Na vapor it is possible to slow light down to an almost complete stop. Inform yourself with the help of some literature. Which effects apart from the extremely flat dispersion relation for large k-vectors contribute to this phenomenon?

8.6. Calculate the dispersion relation $\omega(\boldsymbol{k})$ from (8.7) for vanishing damping. What changes if $\omega_0 = \omega_0(\boldsymbol{k})$ and/or if a small but finite damping γ are introduced?

8.7. Write down the equations of motion of two coupled harmonic oscillators and try to solve them. Find or imagine examples in classical physics and in quantum mechanical systems.

8.8. Which dependence of $\Delta_{\mathrm{LT}}(\boldsymbol{k})$ do you expect for a transition that is dipole forbidden but allowed in quadrupole approximation?

8.9. Can you give an intuitive explanation as to why a photon with spin $\pm\hbar$ can excite a quadrupole transition, e.g., from an atomic s state to a d (or s) state? Hint: Place the atom in the origin of the coordinate system and vary the impact parameter of the photon.

8.10. Sketch the dispersion of a polariton resonance with spatial dispersion and an oscillator strength which increases with $\boldsymbol{k}$. (Assume for simplicity zero damping). Does Δ_{LT} then also depend on $\boldsymbol{k}$?

8.11. Sketch the dispersion of the polariton for two close lying resonances A and B, with and without spatial dispersion for an order of the energies at $\boldsymbol{k}=0$ $\hbar\omega_0^A < \hbar\omega_{\mathrm{L}}^A < \hbar\omega_0^B < \hbar\omega_{\mathrm{L}}^B$. Is it possible for a single orientation of the polarization to have the sequence $\hbar\omega_0^A < \hbar\omega_0^B < \hbar\omega_{\mathrm{L}}^A < \hbar\omega_{\mathrm{L}}^B$?

8.12. Apart from the use of ATR methods see Sect. 21.1.5, it is possible to excite surface polaritons optically if a periodic structure, i.e., a grating, is formed at the interfaces. What is the principle behind this? Compare this with the statements about momentum conservation in Sects. 5.1.3 and 9.2.

References

[02W1] J. Wood, Philos. Mag. **3**, 128 (1902); Phys. Z. **3**, 230 (1902)
[03L1] O. Lummer, E. Pringsheim, Phys. Z. **4**, 430 (1903)
[09M1] L. Pfaundler (ed.), *Müller-Pouillets Lehrbuch der Physik und Meteorologie* 2. Band, 3. Buch, (Vieweg, Braunschweig, 1909), p. 227
[46B1] F. Bloch, Phys. Rev. **70**, 460 (1946)
[58H1] J.J. Hopfield, Phys. Rev. **112**, 1555 (1958)
[62P1] S.J. Pekar, Sov. Phys. Solid State **4**, 953 (1962)
[63H1] J.J. Hopfield, D.G. Thomas, Phys. Rev. **132**, 563 (1963)
[70G1] C.G.B. Garrett, D.E. McCumber, Phys. Rev. A **1**, 305 (1970)
[72B1] A.S. Barker, R. Loudon, Rev. Mod. Phys. **44**, 18 (1972)
[74A1] G.S. Agarwal, Phys. Rev. B **10**, 1447 (1974)
[74O1] A. Otto, Festkörperprobleme/Adv. Solid State Phys. **14**, 1 (1974)
[75C1] R. Claus, L. Merten, J. Brandmüller, *Light Scattering by Phonon Polaritons*. Springer Tracts in Modern Physics, vol. 75 (Springer, Berlin, 1975)
[75L1] J. Lagois, K. Hümmer, Phys. Status Solidi (b) **72**, 393 (1975)
[76P1] R.W. Pohl, *Optik und Atomphysik*, 13th edn. (Springer, Berlin, 1976)
[78B1] I. Broser et al., Phys. Status Solidi (b) **90**, 77 (1978)
[78H1] K. Hümmer, P. Gebhardt, Phys. Status Solidi (b) **85**, 271 (1978)
[78S1] W. Stössel, H.J. Wagner, Phys. Status Solidi (b) **89**, 403 (1978)
[79B1] I. Broser, M. Rosenzweig, Phys. Status Solidi (b) **95**, 141 (1979)
[79M1] Y. Masumoto et al., J. Phys. Soc. Jpn. **47**, 1844 (1979)
[79S1] A. Stahl, Ch. Uhilein, Festkörperprobleme/Adv. Solid State Phys. **XIX**, 159 (1979)
[79U1] R.G. Ulbrich, G.W. Fehrenbach, Phys. Rev. Lett. **43**, 963 (1979)
[80R1] H. Raether, *Excitation of Plasmons and Interband Transitions by Electrons*. Springer Tracts in Modern Physics, vol. 88 (Springer, Berlin, 1980)
[81B1] I. Balslev, Phys. Rev. B **23**, 3977 (1981)
[81I1] T. Itho et al., Solid State Commun. **37**, 925 (1981)
[81L1] J. Lagois, Phys. Rev. B **23**, 5511 (1981)
[81S1] A. Stahl, Phys. Status Solidi (b) **106**, 575 (1981)
[82O1] Y. Onodera, J. Phys. Soc. Jpn. **51**, 2194 (1982)
[82R1] M. Rosenzweig, Excitonische Polaritonen – optische Eigenschaften räumlich dispersiver Medien, Dissertation, Berlin (1982)
[82S1] A. Stahl, I. Balslev, Phys. Status Solidi (b) **111**, 531 (1982); ibid. **113**, 583 (1982)
[82C1] S. Chu, S. Wong, Phys. Rev. Lett **48**, 738 (1982)
[82M1] V.M. Agranovich, A.A. Maradudin (eds.), *Modern Problems in Condensed Matter Sciences* (North Holland, Amsterdam, 1982); *Surface Polaritons*, vol. 1, ed. by V.M. Agranovich, D.L. Mills and *Surface Excitations*, vol. 9, ed. by V.M. Agranovich, R. Landon
[84M1] M. Matsushito, J. Wicksted, H.Z. Cummins, Phys. Rev. B **29**, 3362 (1984)
[84A1] V.M. Agranovich, V.L. Ginzburg, *Crystal Optics with Spatial Dispersion and Excitons*. Springer Series in Solid-State Sciences, vol. 42, 2nd edn. (Springer, Berlin, 1984)
[84H1] P. Halevi, R. Fuchs, J. Phys. C. **17**, 3869, 3889 (1984)
[84K1] C. Klingshirn, *Energy Transfer Processes in Condensed Matter* (1983). NATO ASI Series B, vol. 114 (Plenum Press, New York, 1984), p. 285
[84R1] J.E. Rothenberg, D. Grischkovsky, A.C. Balant, Phys. Rev. Lett. **53**, 552 (1984)
[84S1] T. Shigenari, X.Z. Lu, H.Z. Cummins, Phys. Rev. B **30**, 1962 (1984)
[85H1] B. Hönerlage et al., Phys. Rep. **124**, 161 (1985)
[88R1] H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*. Springer Tracts in Modern Physics, vol. 111 (Springer, Berlin, 1988)
[89R1] V. Ya Reznichenko, M.I. Strashnikova, V.V. Cherny, Phys. Status Solidi (b) **152**, 675 (1989)
[91F1] D. Fröhlich et al., Phys. Rev. Lett. **67**, 2343 (1991)
[91K1] W. Kress, F.W. de Wette (eds.), *Surface Phonons* (Springer, Berlin, 1991)

[91R1] V. Ya Reznichenko, M.I. Strashnikova, V.V. Cherny, Phys. Status Solidi (b) **167**, 311 (1991)
[93H1] H. Haug, S.W. Koch, *Quantum Theory of the Optical and Electronic Properties of Semiconductors*, 2nd edn. (World Scientific, Singapore, 1993)
[93P1] N. Peyghambarian, S.W. Koch, A. Mysyrowicz, *Introduction to Semiconductor Optics* (Prentice Hall, Englewood Cliffs, 1993)
[94R1] G. Raithel, Phys. Bl. **50**, 1149 (1994)
[97N1] S. Nüsse et al., Phys. Rev. B **55**, 4620 (1997)
[97S1] A.J. Sievers, *Spectroscopy and Dynamics of Collective Excitations in Solids* (1995). NATO ASI Series B, vol. 356 (Plenum Press, New York, 1997), p. 227
[97W1] U. Woggon, *Optical Properties of Semiconductor Quantum Dots*. Springer Tracts in Modern Phyiscs, vol. 136 (Springer, Berlin, 1997)
[98B1] R.v. Baltz, C. Klingshirn, *Ultrafast Dynamics of Quantum Systems* (1997). NATO ASI Series B, vol. 372 (Plenum Press, New York, 1998), p. 381
[98B2] G. Benedek, *Ultrafast Dynamics of Quantum Systems* (1997). NATO ASI Series B, vol. 372 (Plenum Press, New York, 1998), p. 295
[98H1] T. Ha et al., Phys. Rev. Lett. **80**, 2093 (1998)
[98K1] C. Klingshirn, *Ultrafast Dynamics of Quantum Systems* (1997). NATO ASI Series B, vol. 372 (Plenum Press, New York, 1998), p. 143
[98J1] A. Jolk, C. Klingshirn, R.v. Baltz, *Ultrafast Dynamics of Quantum Systems* (1997). NATO ASI Series B, vol. 372 (Plenum Press, New York, 1998), p. 397
[99B1] U. van Bürck, Hyperfine Interact. **123/124**, 483 (1999); see also ibid. **125** (1999)
[99C1] T.C. Choy, *Effective Medium Theory - Principles and Applications* (Oxford Science, Oxford, 1999)
[99H1] L.V. Hau et al., Nature **397**, 594 (1999)
[99H2] K. Henneberger, Phys. Rev. Lett. **83**, 1265 (1999)
[99K1] M.M. Kash et al., Phys. Rev. Lett. **82**, 5229 (1999)
[00F1] M. Fleischhauer, M.D. Lukin, Phys. Rev. Lett. **84**, 5094 (2000)
[00K1] C. Klingshirn, Physik in unserer Zeit **31**(4), 144 (2000)
[01K1] O. Kocharovskaya, Y. Rostovtsev, M.O. Scully, Phys. Rev. Lett. **86**, 628 (2001)
[01N1] J.S. Nägerl et al., Phys. Rev. B **63**, 235202 (2001)
[01S1] H.C. Schneider et al., Phys. Rev. B **63**, 045202 (2001)
[02M1] E.A. Muljarov, R. Zimmermann, Phys. Rev. B **66**, 235319 (2002)
[03B1] M. Bajcsy, A.S. Zibrov, M.D. Lukin, Nature **426**, 638 (2003)
[03C1] K. Cho, *Optical Response of Nanostructures: Microscopic, Nonlocal Theory*. Springer Tracts in Solid State Sciences, vol. 139 (Springer, Berlin, 2003)
[04S1] S. Schuhmacher et al., Phys. Rev. B **70**, 235340 (2004)
[05A1] A. André et al., Phys. Rev. Lett. **94**, 063902 (2005)
[05F1] D. Fröhlich et al., Solid State Commun. **134**, 139 (2005)
[05S1] G.V. Smirnov et al., Phys. Rev. A **71**, 023804 (2005)
[06G1] A.D. Greentree et al., Nat. Phys. **2**, 856 (2006)
[06H1] M.J. Hartmann, F.G.S.L. Brandão, M.B. Plenio, Nat. Phys. **2**, 849 (2006); Phys. J. **9**(4), 41 (2010)
[06K1] L. Karpa, M. Weitz, Nat. Phys. **2**, 332 (2006)
[06Z1] F.E. Zimmer et al., Opt. Commun. **264**, 441 (2006)
[07M1] D. Meschede, *Optics, Light and Lasers*, 2nd edn. (Wiley-VCH, Weinheim, 2007)
[07S1] B.E.A. Saleh, M.C. Teich, *Fundamentals of Photonics*, 2nd edn. (Wiley, Hoboken, 2007)
[07S2] V.M. Shalaev, S. Kawata (eds.), *Nanophotonics with Surface Plasmons* (Elsevier B.V., Amsterdam, 2007)
[10O1] J. Otterbach et al., Phys. Rev. Lett. **104**, 033903 (2010)
[10S1] D. Sarid, W.A. Challener, *Modern Introduction to Surface Plasmons: Theory, Mathematica Modeling, and Applications* (Cambridge University Press, Cambridge, 2010)
[14D1] W. Demtröder, *Laser Spectroscopy 1, Basic Principles*, 5th edn. (Springer, Berlin, 2014)
[15O1] L.C. Oliveira, A.M. Nogueira Lima, C. Thirstrup, H.F. Neff, *Surface Plasmon Resonance Sensors: A Materials Guide to Design and Optimization* (Springer, Cham, 2015)

Part II
Phonons

Part II of this book on semiconductor optics is intended to summarize the optical properties of the quanta of lattice vibrations, i.e., phonons. To this end we start with the description of the crystal lattice in real space. The consequences of a periodic arrangement of the crystal atoms on the propagation of waves-like excitations (like phonons) require the consideration of dispersion relations in reciprocal space. We will discuss why such excitations behave like "quasi-particles" and characteristic properties like density of states and statistics.

Of didactic importance is that optical phonons coupled to light behave very much like the model system for polaritons described in Chap. 8. Optical properties like the phonon stop band can thus easily be understood in the polariton picture. We also describe how to study phonons by infrared spectroscopy or light scattering. Phonon polaritons propagating along surfaces will be addresses as well as phonons in disordered materials and in low-dimensional semiconductor systems.

Chapter 9
The Crystal Lattice and Reciprocal Space

Abstract In this chapter we start to discuss topics that are specific to crystalline solids and apply to the large majority of inorganic semiconductors. In particular we lay out the basis for the following chapters where we consider elementary excitations and so-called quasi-particles in semiconductors. These will be needed to describe and understand the linear optical properties. The properties of the quasi-particles themselves are governed by the underlying periodicity of the crystalline lattice. The latter is in particular reflected by the dispersion of the quasi-particles in reciprocal space.

9.1 Description of Periodicity in Real Space

Mostly we will discuss in this book crystalline semiconductors. Crystalline solids have a periodic spatial arrangement of atoms, i.e., they show long-range order. Consequences of deviations from perfect periodicity like in disordered or amorphous systems will be mentioned explicitly.

For *crystal lattices* we can define three non-coplanar elementary translation vectors $\boldsymbol{a}_i (i = 1, 2, 3)$ with the property that if we start at a special atom, e.g., a Ga atom in a GaAs crystal, we reach an identical atom after a spatial shift by a vector $\boldsymbol{R}$ given by

$$\boldsymbol{R} = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3 \tag{9.1}$$

with $n_i = 0, \pm 1, \pm 2, \ldots$

The vector $\boldsymbol{R}$ is called a *translation vector* of the lattice. If we shift the (infinitely large) lattice by $\boldsymbol{R}$ it is transferred onto itself.

The vectors $\boldsymbol{a}$ define a parallelepiped which is called the *unit cell* (see also Figs. 9.1 and 9.2). The whole volume of a crystal is completely filled with identical unit cells. The unit cell and the vectors $\boldsymbol{a}_i$ are called primitive if the unit cell has the minimum possible volume. This definition is not unique as we realize for the example of a two-dimensional cubic lattice in Fig. 9.1, where we show a non-primitive unit cell and two primitive ones. By convention, a special primitive unit cell is agreed upon. In the case described here the one defined by $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$.

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_9

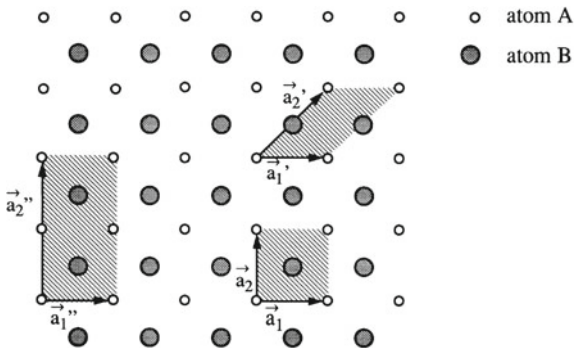


Fig. 9.1 Two primitive (*r.h.s*) and a non-primitive (*l.h.s.*) unit cells in a two-dimensional cubic lattice with a basis consisting of two different atoms per primitive unit cell

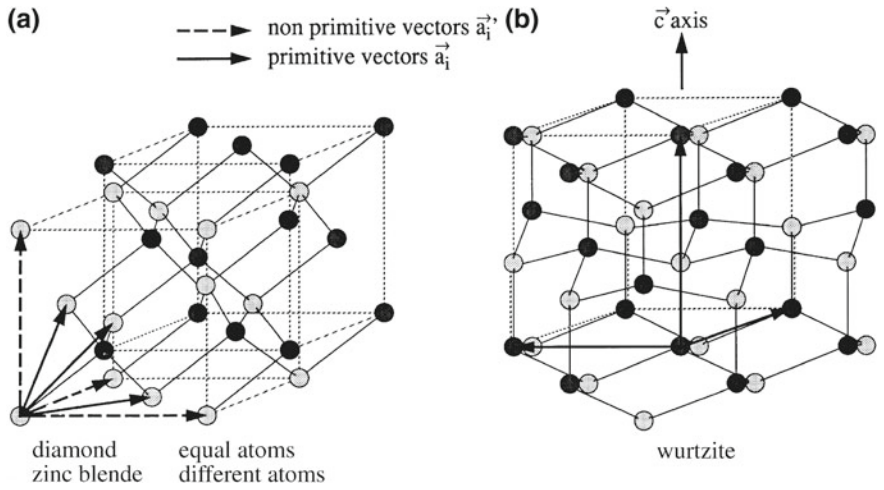


Fig. 9.2 The unit cells for the diamond and zinc-blende type crystal structures (**a**) and of the wurtzite crystal structure (**b**) (reprinted with permission from [82L1] ©1982–2001 by Springer Nature)

The vectors $\boldsymbol{R}$ evidently form for an infinite crystal an Abelian group which is called the *translational group* (see Appendix B). The positions of the atoms in the unit cell are given by the so-called *basis*. In Fig. 9.1 the basis consists of two atoms, one atom A at (0, 0) and one atom B at $(1/2\boldsymbol{a}_1, 1/2\boldsymbol{a}_2)$. The translation vectors $\boldsymbol{a}_i$ define an abstract, translation invariant point lattice while the basis gives the information where the atoms are really located in the primitive unit cell. The lattice and the basis define together the *crystal structure*.

Apart from the translational group there is another type of symmetry operations which transforms the lattice into itself, but for which at least one point is kept fixed. These symmetry operations also form a group which is called the *point group*. The

elements of this group are for example reflections at mirror planes, rotations around axes with two-, three-, four- or six-fold symmetry or the inversion through the origin. Furthermore there may be screw axes or glide planes, which combine either a rotation axis or a mirror plane with a translation by a rational fraction of the $\boldsymbol{a}_i$.

The abstract translation lattices can be grouped into 14 *Bravais lattices* namely one triclinic, two monoclinic, four orthorhombic, two tetragonal, three cubic, one trigonal (rhombohedral) and one hexagonal lattice. If we include the positions of the atoms and the translational invariance we find from all possible combinations of the symmetry operations which transform an infinite crystal into itself a total of 230 so-called *space groups*, out of which 73 can be written as a product of the translation group and the point group. For more details see Appendix B and the references therein.

The most important point groups for semiconductors are O_h (realized e.g., in the crystal structures of diamond, Si, Ge, Cu_2O or NaCl), T_d (realized e.g., in the zincblende type crystal structure of ZnS, ZnSe, GaAs, InP, CuCl or AgBr) and C_{6v} (realized e.g., in the wurtzite type crystal structure of ZnS, ZnO, CdS or GaN).

We show in Fig. 9.2 the crystal structures of diamond, zincblende and wurtzite. The diamond crystal structure consist of C atoms occupying the lattice points of two face-centered cubic (fcc) lattices shifted by $1/4$ of the space diagonal of the cubic unit cell. For zincblende the same principle applies, however, one of the two sublattices is occupied by atoms A, the other by B. This situation is depicted in Fig. 9.2a. If the grey and the black "atoms" are identical, then the diamond crystal structure results. In Fig. 9.2a we show three non-primitive translation vectors $\boldsymbol{a}'_i$ by the dashed arrows, which just define one of the fcc cubes. By solid arrows we show the primitive translation vectors $\boldsymbol{a}_i$, which start at a corner of the cube and point to the three nearest identical atoms, which are situated in the middle of the three faces intersecting in the corner. The wurtzite crystal structure is hexagonal with a polar crystallographic $\boldsymbol{c}$-axis. In this case, one has two primitive translation vectors frequently called $\boldsymbol{a}$ and $\boldsymbol{b}$ of equal length which enclose an angle of 120° and a third one (called $\boldsymbol{c}$) of different length and perpendicular to the plane spanned by $\boldsymbol{a}$ and $\boldsymbol{b}$. Usually the z-axis is chosen parallel to $\boldsymbol{c}$. In all three cases, one atom is surrounded by a tetrahedron of its four nearest neighbors. The difference between zincblende and wurtzite structures is in the positions of the next nearest neighbors only. Therefore several of the above mentioned compound semiconductors can crystallize in both structures like ZnS (which is notorious for these two polytypes), CdS or GaN.

The chemical binding of the semiconductors is covalent for the elements (C, Si, Ge) with sp^3 hybridization, and acquires an increasing and finally dominant ionic admixture when going to the III–V, II^b–VI and I^b–VII compounds.

9.2 Reciprocal Space and Quasi-momentum

Now we want to introduce the so-called *reciprocal lattice*. It is defined by its elementary translation vectors $\boldsymbol{b}_i$ in the same way as the lattice in real space. The $\boldsymbol{b}_i$ are given by:

$$\boldsymbol{b}_1 = \frac{2\pi}{V_{\mathrm{uc}}} \boldsymbol{a}_2 \times \boldsymbol{a}_3 \tag{9.2}$$

and cyclic permutations of the indices, V_{uc} is the volume of the unit cell in real space given by

$$V_{\mathrm{uc}} = \boldsymbol{a}_1(\boldsymbol{a}_2 \times \boldsymbol{a}_3). \tag{9.3}$$

A general translation vector of the reciprocal lattice is usually called $\boldsymbol{G}$

$$\boldsymbol{G} = l_1\boldsymbol{b}_1 + l_2\boldsymbol{b}_2 + l_3\boldsymbol{b}_3 \quad l_i = 0, \pm 1, \pm 2, \ldots \quad i = 1, 2, 3\ . \tag{9.4}$$

Without trying to be complete, we give some properties of the reciprocal lattice and its connections with the real one. Every periodic function in real space which is sufficiently smooth and has a periodicity given by $f(\boldsymbol{r} + \boldsymbol{R}) = f(\boldsymbol{r})$ and $\boldsymbol{R}$ defined by (9.1) can be written as a Fourier series summing over all vectors of the reciprocal lattice

$$f(\boldsymbol{r}) = \sum_{\boldsymbol{G}} f_{\boldsymbol{G}} e^{\mathrm{i}\boldsymbol{G}\boldsymbol{r}} \tag{9.5a}$$

with

$$f_{\boldsymbol{G}} = \boldsymbol{V}_{\mathrm{uc}}^{-1} \int_{\mathrm{uc}} f(\boldsymbol{r}) e^{-\mathrm{i}\boldsymbol{G}\boldsymbol{r}} \mathrm{d}\tau\ . \tag{9.5b}$$

The scalar product of $\boldsymbol{R}$ and $\boldsymbol{G}$ always fulfills

$$\boldsymbol{R} \cdot \boldsymbol{G} = 2\pi m; \quad m = 0, \pm 1, \pm 2, \ldots \tag{9.6}$$

As a consequence, we can choose to describe effects occurring in periodic lattices in real space or in reciprocal space. The latter is the appropriate space for wave vectors $\boldsymbol{k}$ or (quasi-)momenta $\hbar\boldsymbol{k}$. The transformation from one space into the other is given by the three-dimensional Fourier series of (9.5).

A crystalline lattice no longer has invariance with respect to infinitesimal translations in space (Sect. 5.1.3) but only invariance with respect to translations by integer multiples of $\boldsymbol{a}_i$. The conservation law for the momentum $\hbar\boldsymbol{k}$ which follows from an invariance with respect to infinitesimal translations according to Noether's theorem (5.14b) is modified for a periodic lattice such that $\hbar\boldsymbol{k}$ is conserved only to within integer multiples of the $\boldsymbol{b}_i$, i.e., we can add to a given $\boldsymbol{k}$-vector an arbitrary vector from the reciprocal lattice $\boldsymbol{G}$:

$$\boldsymbol{k} \leftrightarrows \boldsymbol{k} + \boldsymbol{G}\ . \tag{9.7}$$

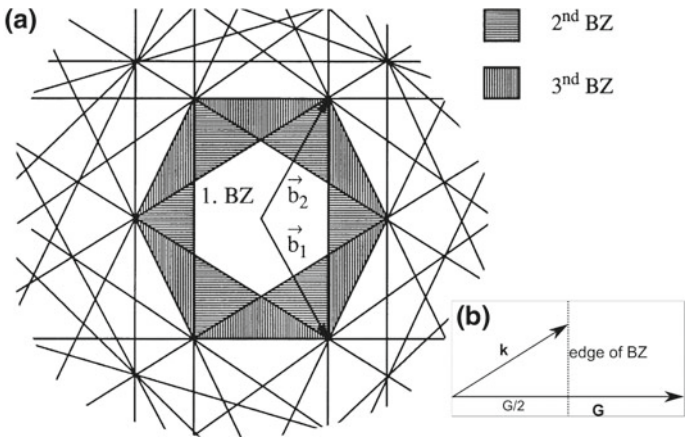


Fig. 9.3 The first Brillouin zones of a two-dimensional, hexagonal lattice **(a)** and illustration of a wavevector pointing to the edge of the Brillouin zone fulfilling the projection condition $\boldsymbol{k} \cdot \boldsymbol{G}/2 = |\boldsymbol{G}|^2/4$ **(b)**

From (9.7) it is clear that we do not have to consider the whole $\boldsymbol{k}$-space, but discussion of physics can be restricted to a "unit-cell" defined by the vectors $\boldsymbol{b}_i$. Every $\boldsymbol{k}$-vector that is outside the unit cell can be shifted inside the unit cell by adding an appropriate $\boldsymbol{G}$. Usually one does not work in reciprocal space with unit cells defined as in Fig. 9.1 or Fig. 9.2, but uses a different construction illustrated in Fig. 9.3a for the two-dimensional case. One constructs the planes perpendicularly bisecting the lines connecting one point of the reciprocal lattice, which is chosen as the origin, with all others. The figure thereby enclosed around the origin is called the first *Brillouin zone*; the equivalent pieces which are next together form the second Brillouin zone, and so on. All Brillouin zones have equal area or volume in two or three dimensions, respectively. All higher Brillouin zones can be shifted into the first one by adding appropriate $\boldsymbol{G}$ vectors.

Equation (9.7) is a very important statement which forms, together with energy conservation, the basis for the theoretical description of diffraction of X-rays, electrons or neutrons from a periodic lattice. More general, propagation of any wave (including (quasi-)particles like electrons, phonon or excitons whose propagation is described in terms of wave packets) in the periodic lattice potential leads to scattering from all the atoms and thus to interference of all partial waves. We have discussed this already for the propagation of polaritons in dense media, where constructive interference of the scattered partial waves only occurs in forward direction (see Sect. 8.3.2). A special situation occurs if the wave vector of the wave fulfills the condition $\boldsymbol{k} \cdot \boldsymbol{G}/2 = |\boldsymbol{G}|^2/4$, i.e., the projection of the vector $\boldsymbol{k}$ onto a reciprocal lattice vector $\boldsymbol{G}$ amounts to half of the length of $|\boldsymbol{G}|$ (see Fig. 9.3b). Mathematically this is exactly the instruction how to construct the Brillouin zone given above.

This special condition is actually equivalent to the Laue condition for diffraction (see e.g., [16A1]) which means that all waves with a wave-vector whose tip ends

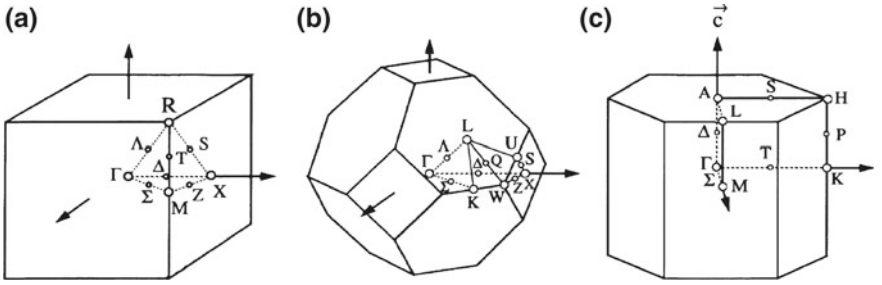


Fig. 9.4 The first Brillouin zones for a simple cubic lattice (**a**), of the diamond and zinc-blende-type structures (point groups 0_h, and T_d, respectively) (**b**) and of the hexagonal wurtzite-type structure (C_{6v}) (**c**). The notation of points and of directions of high symmetry are indicated (reprinted with permission from [82L1] ©1982–2001 by Springer Nature)

on the rim of the Brillouin zone fulfills the Laue condition. This has distinct consequences for the wave propagation. It results in a constructive interference of the forward wave and all waves scattered backward from the periodic atom arrangement leading to a standing wave. Waves with this specific wave vectors cannot propagate and have a vanishing group velocity. This is reflected in horizontal dispersion relations for phonons, electrons etc. at the edge of the Brillouin zone. We will come back to this property of quasi-particle propagation when describing the respective dispersion relations (see e.g., Sects. 11.1, 14.2).

The Brillouin zones also form a type of elementary cell, but constructed according to Fig. 9.3 and not according to Fig. 9.1. The unit cell constructed in real space according to Fig. 9.3 is known as *Wigner–Seitz cell*. The first Brillouin zone is thus termed the Wigner–Seitz cell of the reciprocal lattice. The names of points and lines of high symmetry in the first Brillouin zone are indicated in Fig. 9.4.

For a simple cubic lattice with

$$\boldsymbol{a}_1 = (a, 0, 0), \quad \boldsymbol{a}_2 = (0, a, 0), \quad \boldsymbol{a}_3 = (0, 0, a) \,, \tag{9.8}$$

the $\boldsymbol{b}_i$ are also orthogonal and read

$$\boldsymbol{b}_1 = \left[\frac{2\pi}{a}, 0, 0\right], \; \boldsymbol{b}_2 = \left[0, \frac{2\pi}{a}, 0\right], \; \boldsymbol{b}_3 = \left[0, 0, \frac{2\pi}{a}\right] \,, \tag{9.9}$$

and the first Brillouin zone is a cube whose extension in all three directions is given by

$$-\frac{\pi}{a} \leq k_i \leq +\frac{\pi}{a}, \quad i = x, y, z \,. \tag{9.10}$$

In many of the following figures and equations, we give for simplicity the values of (9.10) for the borders of the first Brillouin zone, though most solids including

semiconductors are not crystallizing in a simple cubic lattice but frequently in the structures discussed above in Fig. 7.2 or others.

In Fig. 9.4 we give the first Brillouin zones for this simple cubic lattice and for the point groups T_d, O_h and C_{6v} using the primitive unit cell including the notation for some special points and directions in reciprocal space. The center of the first Brillouin zone $\boldsymbol{k} = (0, 0, 0)$ is always called the Γ-*point*, other points of high symmetry are labeled by capital Latin letters and directions of high symmetry by capital Greek letters. As an example: in T_d symmetry, when leaving the Γ-point in the Σ-direction one arrives in at the point K at the border of the first Brillouin zone.

The quantity $\hbar\boldsymbol{k}$ of an excitation or quasi-particle (phonon, exciton, electron, hole ...) in a periodic lattice is usually called *quasi-momentum*. This notation stresses the difference with respect to the momentum $\boldsymbol{p} = \hbar\boldsymbol{k}$ of a free particle, e.g., a photon or an electron in vacuum. For the case of a periodic potential as in a crystal the Hamiltonian has not the full invariance with respect to translations in space. Thus its eigenfunctions (e.g., Bloch functions for the case of electrons) are not simultaneously eigenfunctions of the momentum operator. Applying the momentum operator to the eigenfunction does not lead to a momentum $\boldsymbol{p} = \hbar\boldsymbol{k}$ as in the case of the plane-wave eigenfunction of a free particle. The quasi-momentum is thus not a momentum but a quantum number of the quasi-particle which actually can be used similar to a momentum to deduce the reaction of the quasi-particle when forces are applied. Further a conservation law for quasi-momenta applies e.g., for scattering of quasi-particles in periodic potentials. But, in contrast to scattering of free particles, arbitrary reciprocal lattice vectors $\boldsymbol{G}$ can be added when balancing the quasi-momenta.

Actually it is possible to make a transition from the quasi-particle to the free-particle case: if the lattice constant a goes to zero, the system regains translational invariance with respect to infinitesimally small shifts in real space. On the other hand, the vectors $\boldsymbol{b}_i$ defined in (9.9) go to infinity in this limit and the first Brillouin zone fills the whole $\boldsymbol{k}$-space, so that reciprocal lattice vectors become physically meaningless.

We will come back to the a discussion of the term quasi-momentum e.g., in Sect. 11.1. More details are also given in books on solid state physics like [16A1].

Problems

9.1. Inspect or build some lattice and crystal models to become familiar with the topics presented in Sect. 9.1.

9.2. Show that the reciprocal lattice of a face-centered cubic (fcc) lattice is a body centered cubic (bcc) lattice and vice versa.

9.3. Calculate and draw the primitive unit cell and the Wigner–Seitz cell in real space and the first three Brillouin zones in reciprocal space for a simple cubic and a hexagonal two-dimensional lattice.

9.4. Inspect a model of a cubic crystal (e.g., zinc-blende). Find the non primitive cubic unit cell and the primitive one. Explain qualitatively that such a crystal should be optically isotropic for light propagating ($\boldsymbol{k} \neq 0$!) e.g., in the directions (100) or (111) but not in (110).

References

[82L1] O. Madelung, U. Rössler (ed.), *Landolt–Börnstein*. New Series, Group III, vol. 17a to i, 22a and b, 41A to D (Springer, Berlin, 1982–2001)

[16A1] N.W. Ashcroft, N.D. Mermin, D. Wei, *Solid State Physics: Revised Edition* (Cengage Learning Asia, Singapore, 2016)

Chapter 10
Vibrations of Atoms in a Crystal Lattice

Abstract In this chapter we treat the vibrations of atoms in a crystalline lattice in the way introduced in Frenkel (Phys Rev 37:1276, 1931) [31F1] and adopted in many textbooks (e.g., Ashcroft et al. in Solid State Physics. Cengage Learning Asia, Singapore, 2016 [16A1]). We will start with a homogeneous string, proceed to monoatomic and diatomic chains and finally arrive at the three-dimensional arrangement found in solids. The resulting quanta of these lattice vibrations, the phonons, will be the focus of Chap. 11.

10.1 Vibrations of a String

Let us first consider a quasi *one-dimensional string*, as shown schematically in Fig. 10.1. Two types of waves can propagate along such a string, transverse and longitudinal ones. The direction of the elongation is either perpendicular to the direction of propagation, i.e., in the x–y plane, or parallel to it, i.e., in the z-direction, respectively. We start with the latter case. The mass density of the string is ρ, its cross-section A, and the elongation of an infinitesimally small piece of length dz and mass dm of the string at z from its equilibrium position is $u(z)$. Then, Newton's equation of motion reads

$$\mathrm{d}m \frac{\partial^2 u}{\partial t^2} = \rho A \cdot \mathrm{d}z \cdot \frac{\partial^2 u}{\partial t^2} = F \ . \tag{10.1}$$

The force F is connected to the elasticity modulus E via

$$F = A \cdot E \frac{\partial^2 u}{\partial z^2} \ . \tag{10.2}$$

The appearance of the second derivative in (10.2) might be surprising at first, bearing in mind Hooke's law. However, we must consider that the stress σ is indeed given by

$$\sigma(z) = E \frac{\partial u}{\partial z} \ . \tag{10.3}$$

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_10

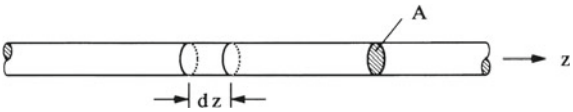


Fig. 10.1 A piece of a string as a model to explain the derivation of (10.4)

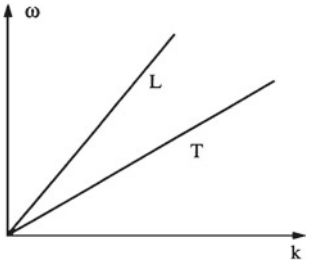


Fig. 10.2 The dispersion relation of waves on a homogeneous string

If the stress is the same on both sides of the infinitesimal element of length $\mathrm{d}z$, the resulting forces at z and $z + \mathrm{d}z$ compensate each other to zero. The restoring force F is therefore given by $\mathrm{d}\sigma/\mathrm{d}z$ leading to (10.2).

Putting (10.1) and (10.2) together leads to the standard harmonic wave equation

$$\rho \frac{\partial^2 u}{\partial t^2} = E \frac{\partial^2 u}{\partial z^2} \,. \tag{10.4}$$

With the ansatz

$$u(z,t) = u_0 \exp\left[\mathrm{i}(kz - \omega t)\right] \tag{10.5}$$

for a plane wave we find the dispersion relation for longitudinal waves

$$\omega_{\mathrm{L}} = (E/\rho)^{1/2} k \,. \tag{10.6}$$

This is a linear relation as shown in Fig. 10.2. Consequently phase and group velocity are constant and equal, namely, with (2.18):

$$v_{\mathrm{ph}}^{\mathrm{L}} = v_{\mathrm{g}}^{\mathrm{L}} = (E/\rho)^{1/2} \,. \tag{10.7}$$

For the two degenerate, transverse waves we find in a similar way

$$\omega_{\mathrm{T}} = (G/\rho)^{1/2} k \tag{10.8}$$

or

$$v_{\mathrm{ph}}^{\mathrm{T}} = v_{\mathrm{g}}^{\mathrm{T}} = (G/\rho)^{1/2} \,. \tag{10.9}$$

where G is the *shear* or *torsion modulus*.

Since it is known from the theory of elasticity that

$$G \leq E \ , \tag{10.10}$$

we find

$$v_{\mathrm{ph}}^{\mathrm{T}} \leq v_{\mathrm{ph}}^{\mathrm{L}} \ , \tag{10.11}$$

a result comparable to (7.28).

10.2 Vibrations of Linear Chains

So far we considered a homogeneous string, neglecting the fact that a solid is made up from individual atoms. Therefore the above approximation can only be valid for wavelengths much larger than the lattice constant or for wave vectors close to the center of the first Brillouin zone, i.e., $\lambda \gg a$ or $k \ll \pi/a$.

For shorter wavelengths we have to consider the atomic structure of solids. The interaction potential between neighboring atoms as a function of the lattice constant a looks qualitatively like Fig. 10.3 for all types of binding, e.g., covalent, ionic, metallic or binding by van der Waals forces. For sufficiently large lattice constants there is no interaction between the atoms, i.e., $V = 0$; for decreasing lattice constant a regime with attractive potential (without which there would be no solids) follows; finally there is a steep repulsive increase due to Pauli's exclusion principle when the filled inner shells of neighboring atoms start to overlap. Different analytic approximations to $V(a)$ are known, e.g., the *Born–Mayer* or *Lennard–Jones potentials*. These details have at present no relevance for us. We note that a crystal acquires a state close to the energetic minimum, i.e., to the equilibrium position a_0. In the vicinity of a_0 the potential $V(a)$ can be approximated by a parabola, that is by a *harmonic potential*, shown by the dashed line. It is at least qualitatively clear that this harmonic approximation is valid only very close to a_0. For larger deviations from a_0 significant anharmonicities (i.e., deviations from the harmonic potential) have to be expected. The anharmonicities are characteristic for lattice vibrations and manifest themselves, among other things, in the thermal expansion of solids and in phonon–phonon interaction.

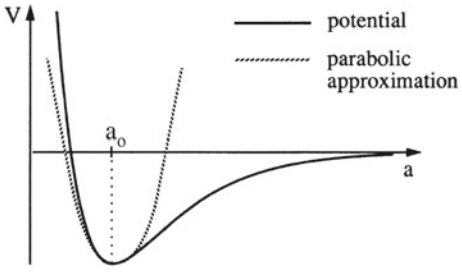


Fig. 10.3 Sketch of the potential V between neighboring atoms in a solid as a function of the lattice constant a

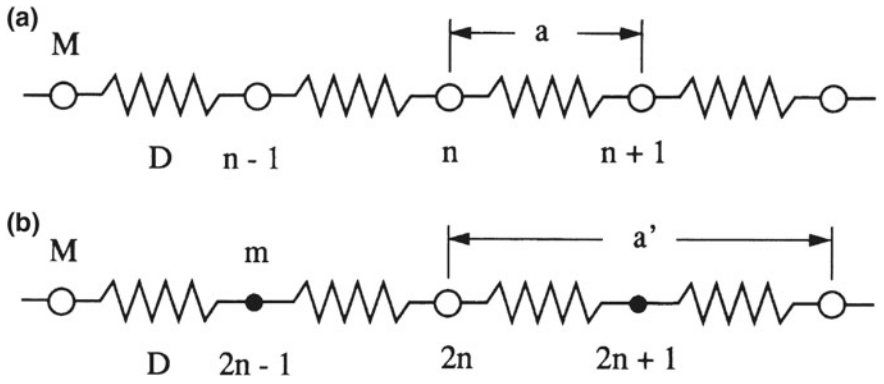


Fig. 10.4 The linear chain model for the cases of one atom per unit cell (**a**) and two (different) atoms per unit cell (**b**)

For the moment, however, the harmonic approximation is adequate and we consider a linear model solid in which every atom with mass M is connected to its neighbors by a "spring" with a force constant D, representing the harmonic potential. This leads to the *linear-chain model* of Fig. 10.4a in which we indicate also the lattice constant a. Evidently we have a basis consisting of one atom per primitive unit cell.

At this point it is important to stress the difference between the models of Figs. 7.1 and 8.8 on one side and of Fig. 10.4 on the other side. In the first case we had independent oscillators attached to their (lattice-)sites and a weak coupling between them was introduced only to simulate spatial dispersion. In the current discussion the coupling springs from the sole forces acting on the atoms.

We introduce now the displacement u_n of atom n from its equilibrium position and obtain its equation of motion, again for the longitudinal mode

$$M\frac{\partial^2 u_n}{\partial t^2} = D\left[(u_{n+1} - u_n) - (u_n - u_{n-1})\right] = D(u_{n+1} - 2u_n + u_{n-1})\ . \quad (10.12)$$

Instead of the second differential quotient in the homogeneous approximation we are now left with a second order difference equation.

As a solution of (10.12) we again choose a plane wave which reads in this discrete case

$$u_n(t) = u_{n,0}\exp\left[\mathrm{i}(kna - \omega t)\right] \quad (10.13a)$$

$$u_{n\pm1}(t) = u_{n\pm1,0}\exp\left\{\mathrm{i}\left[k(n \pm 1)a - \omega t\right]\right\}\ . \quad (10.13b)$$

For a plane wave the amplitudes of the various atoms are equal, i.e.,

$$u_{n,0} = u_{n+1,0} = u_{n-1,0}\ . \quad (10.13c)$$

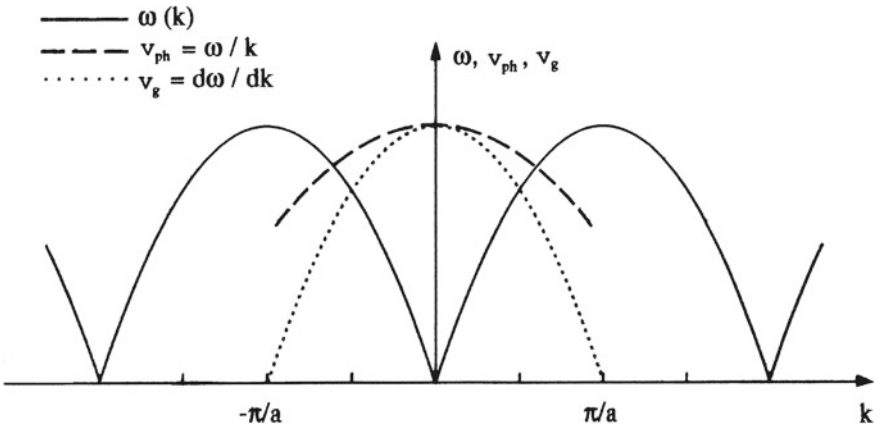


Fig. 10.5 The dispersion relation and the phase and group velocities of the longitudinal vibrations of the *monatomic chain* of Fig. 10.4a

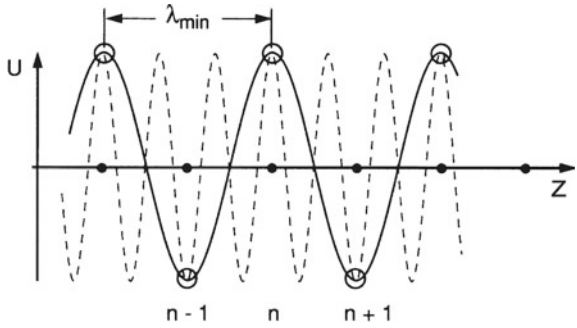


Fig. 10.6 A vibration of the linear chain of Fig. 10.4a with a wave vector inside and outside the first Brillouin zone

Inserting (10.13a), (10.13b) and (10.13c) into (10.12) gives

$$-M\omega^2 = D\left(e^{-\mathrm{i}ka} - 2 + e^{\mathrm{i}ka}\right) = -2D\,(1 - \cos ka) \tag{10.14}$$

or

$$\omega = \left(\frac{4D}{M}\right)^{1/2} \left|\sin\frac{ka}{2}\right| . \tag{10.15}$$

The dispersion relation according to (10.15) is shown in Fig. 10.5 together with the phase and group velocities in the first Brillouin zone.

As can be expected from the discussion in Sect. 9.2, the dispersion relation outside the first Brillouin zone is just a repetition of what is inside, or, in other words, the branches inside and outside can be shifted on top of each other by adding or subtracting reciprocal lattice vectors $l \cdot 2\pi/a$, where l is a positive or negative integer.

The fact that there is no new physics outside the first Brillouin zone can be easily understood for the case of lattice vibrations from Fig. 10.6. If adjacent atoms

are displaced in anti-phase (solid line), the chain acquires the shortest physically meaningful wavelength $\lambda_{\min}$ and the largest wave vector by

$$\lambda_{\min} = 2a \quad \rightarrow \quad k_{\max} = \frac{2\pi}{\lambda_{\min}} = \frac{\pi}{a} \,. \tag{10.16}$$

Of course we can define a shorter wavelength as indicated by the dashed line in Fig. 10.6 resulting for this example in $k = 3\pi/a$. But this case is physically meaningless since there are no atoms at the positions between $z = na$ and $z = (n+1)a$. On the other hand, $k = 3\pi/a$ corresponds to π/a by just adding $G = -2\pi/a$.

We now inspect the phase velocity $v_{\rm ph}$ and group velocity $v_{\rm s}$ in the first Brillouin zone. Except at $k = 0$ the two quantities are no longer equal and change with k. Consequently a wave packet will broaden during propagation, i.e., it will show *dispersion* in the sense discussed in Chap. 8. The fact that the dispersion relation is horizontal at the border of the first Brillouin zone in the direction normal to this border is generally the case. We have illustrated this property in Sect. 9.2 as resulting from the fact that wave-vectors ending on the edge of the Brillouin zone fulfill the Laue condition and thus the related waves are standing waves with vanishing group velocity.

Since most semiconductors have more than one atom per primitive unit cell, we address this situation with the help of Fig. 10.4b where the basis evidently consists of two atoms with masses M and m. The lattice constant is now a'. Using the nomenclature of Fig. 10.4b we obtain the following equations of motion in analogy to (10.12):

$$M \frac{\partial^2 u_{2n}}{\partial t^2} = D\,(u_{2n+1} - 2u_{2n} + u_{2n-1}) \,, \tag{10.17a}$$

$$m \frac{\partial^2 u_{2n+1}}{\partial t^2} = D\,(u_{2n+2} - 2u_{2n+1} + u_{2n}) \tag{10.17b}$$

where we differentiate between the two different types of atoms by using even $(2n)$ and odd $(2n \pm 1)$ indices. Using again the ansatz

$$u_{2n}(t) = u_{2n,0} \exp\left[\mathrm{i}(2nak - \omega t)\right] \,, \tag{10.18a}$$

$$u_{2n+1}(t) = u_{2n+1,0} \exp\left\{\mathrm{i}\left[(2n+1)ak - \omega t\right]\right\} \,, \tag{10.18b}$$

and noting again that the amplitudes of equal masses are equal

$$u_{2n,0} = u_{2n+2,0} = A_{\rm M} \,, \tag{10.19a}$$

$$u_{2n+1,0} = u_{2n-1,0} = A_{\rm m} \,, \tag{10.19b}$$

we get the following set of linear equations:

$$(2D - \omega^2 M) A_{\rm M} - 2D\cos(ka) A_{\rm m} = 0 \,, \tag{10.20a}$$

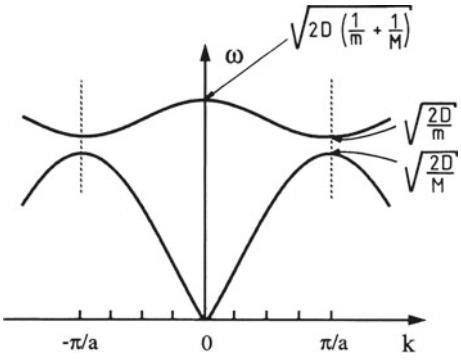


Fig. 10.7 The dispersion relation for the diatomic chain

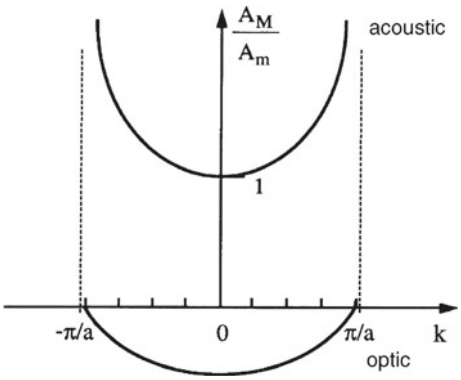


Fig. 10.8 The ratio of the amplitudes A_M/A_m of the two different masses for the two dispersion branches of Fig. 10.7 as a function of k

$$-2D\cos(ka)A_M + (2D - \omega^2 m)A_m = 0\ . \tag{10.20b}$$

These have a non-trivial solution (i.e., one different from $A_M = A_m = 0$) only if the determinant of the coefficients vanishes. The dispersion relation resulting from the corresponding secular equation reads:

$$\omega^2 = D\left(\frac{1}{m}+\frac{1}{M}\right) \pm D\left[\left(\frac{1}{m}+\frac{1}{M}\right)^2 - \frac{4}{Mm}\sin^2\frac{ka'}{2}\right]^{1/2}\ . \tag{10.21}$$

The dispersion relation has now two branches, as shown in Fig. 10.7, where we give also the values at some special points. The lower branch is usually called the *acoustic branch* since sound waves propagate according to its modes. The upper branch is called *optical branch*, for reasons given below.

We can enter the solution (10.21) into (10.20) and calculate the ratio A_M/A_m. The procedure is straightforward but lengthy and so we present the result only graphically in Fig. 10.8 and discuss it in connection with Fig. 10.9. For the acoustic branch, the two different atoms are displaced in the same direction, Fig. 10.9a. For very long wavelengths (i.e., $k \simeq 0$) the amplitudes are equal. Actually the case $k = 0$

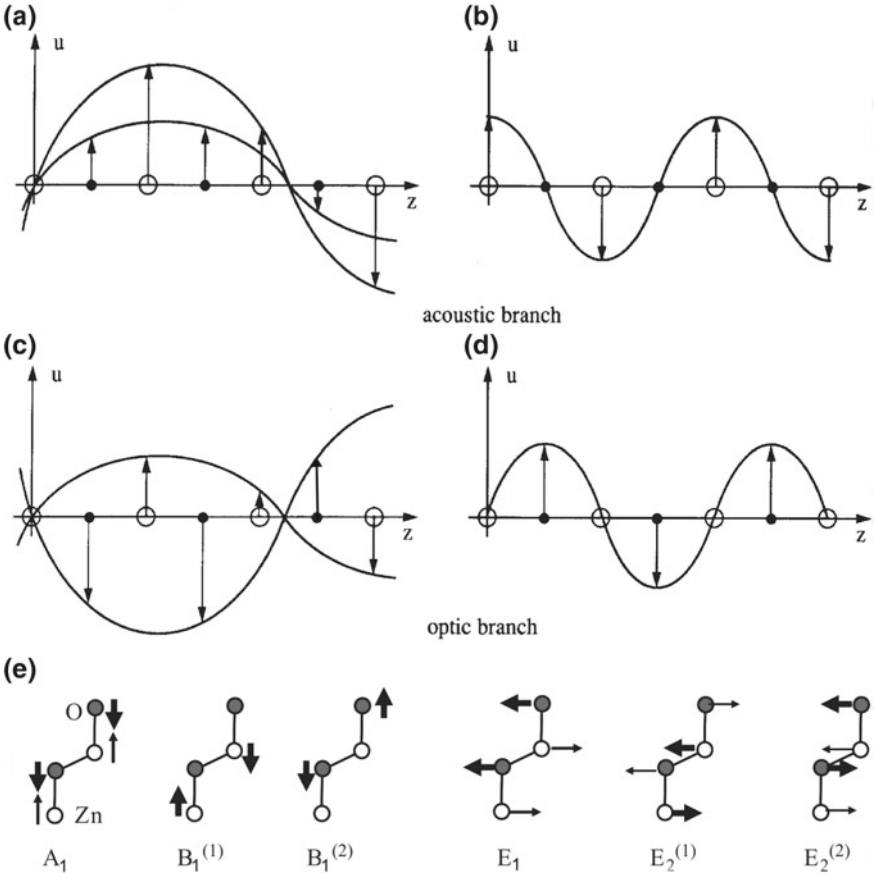


Fig. 10.9 Sketch of the (longitudinal) displacements of the atoms on the acoustic (**a**, **b**) and the optical branches for two different wave vectors (**c**, **d**) and the normal—or eigenmodes of the optical phonons in the wurtzite-type lattices (here for ZnO) (**e**) (adapted with permission from [10G1] ©2010 by Springer Nature)

corresponds to a simple displacement of the whole crystal. For increasing k the amplitude of the heavy mass M gets larger than that of m for $M > m$, and, at the boundary of the first Brillouin zone, only the heavy masses oscillate (Fig. 10.9b), resulting in an eigenfrequency $(2D/M)^{1/2}$ as indicated in Fig. 10.7.

For the optical branch, the two atoms are displaced in anti-phase (Fig. 10.9c). At the boundary of the first Brillouin zone only the light masses m oscillate (Fig. 10.9d).

If the two different atoms carry an electric charge, i.e., if the binding is at least partly ionic, then the oscillation according to Fig. 10.9c is connected with an oscillating electric dipole. This allows a coupling to the electromagnetic light field at least for the transverse eigenmodes and this is why these oscillations are called "optical" modes. We come back to this aspect in Chap. 11.

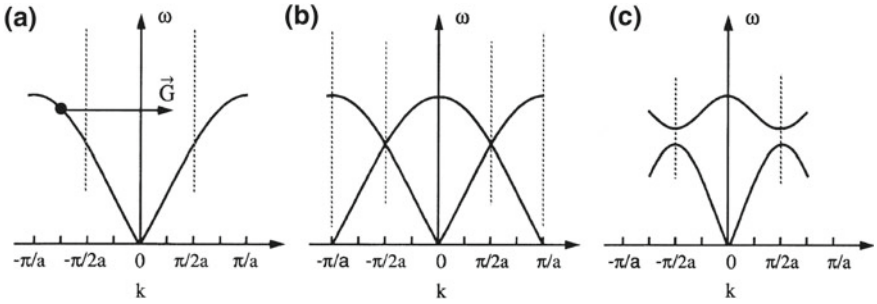


Fig. 10.10 The transition from the monatomic (**a**) to the diatomic (**c**) chain

Here we have discussed in detail only the longitudinal modes, but it is obvious that for both, the monatomic and the diatomic chain, for every wave vector and every branch two (degenerate) transverse oscillations also exist. In this case, the displacements of the atoms are not along the chain but perpendicular to it.

The dispersion relation of Fig. 10.7 can also be deduced in a different way starting from the one of Fig. 10.5, i.e., from the monatomic chain which we repeat in Fig. 10.10a. Now we imagine that we paint the atoms of the monatomic chain alternating in two different colors, but without changing their physical properties. As a consequence we have increased the length of the primitive unit cell by a factor of two and the new lattice constant a' is given by $a' = 2a$. This in turn reduces the length of the first Brillouin zone by one-half as shown in Fig. 10.10a, b. Consequently, we can shift the outer parts of the dispersion relation by vectors $\boldsymbol{G}$ of the new reciprocal lattice into the first Brillouin zone. This situation is shown in Fig. 10.10b. Since the atoms are still identical, the two branches cross at the border of the first Brillouin zone. If we now also introduce differences in the physical properties of the atoms, for example giving them different masses, then we end up with the situation of Fig. 10.10c which is identical to Fig. 10.7. We shall use this set of arguments again in connection with superlattices later on in Sect. 17.2.

10.3 Vibrations in Three-Dimensional Systems

We now want to extend the discussion of atom vibrations to the case of three-dimensional systems. If we consider a three-dimensional crystal, not too much changes occur with respect to the chain model of Sect. 10.2, at least for the dispersion relations. The atoms are connected with nearest (and possibly next-nearest) neighbors in the three-dimensional crystal structure. The set of equations analogous to (10.12) or (10.17) will become correspondingly more complex, but the result will be qualitatively the same. There are still the three acoustic branches and, in addition, optical ones if there is more than one atom in the primitive unit cell. There are

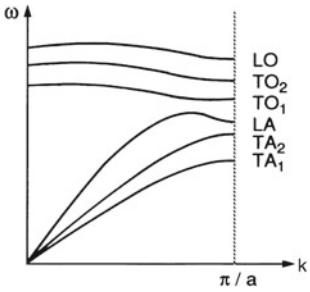


Fig. 10.11 Sketch of the dispersion relation of lattice vibrations for a three-dimensional anisotropic crystal with partly ionic binding and two atoms per primitive unit cell

always three acoustic branches, namely one longitudinal and two transverse ones for every $\boldsymbol{k}$-vector and $3s-3$ optical ones, where s is the number of atoms per primitive unit cell.

If the crystal is anisotropic the dispersion relations will be different for different directions in the Brillouin zone and the degeneracy between the two transverse modes for every $\boldsymbol{k}$ may be lifted. We show such a situation schematically in Fig. 10.11.

Problem

10.1. Find in [82L1] or in the internet the phonon dispersion relations for typical semiconductors like Ge, Si, GaAs and ZnO.

(a) Compare the dispersion curves and give qualitative arguments for the differences.
(b) Determine the energy $\hbar\omega$ of the optical phonons at k=0 and compare it to the thermal energy $k_{\mathrm{B}}T \approx 26$ meV at $T = 300$ K. What do you conclude for the probability that these lattice vibrations are thermally excited at room temperature?
(c) Explain the degree of degeneracy of the TA and TO modes, respectively.
(d) Compare the LT splittings of the optical modes and explain the differences.
(e) Discuss the splitting or degeneracy at the edge of the first Brillouin zone of the LA and LO modes of each semiconductor in comparison to Fig. 10.7.

References

[31F1] J. Frenkel, Phys. Rev. **37**, 1276 (1931)
[82L1] O. Madelung, U. Rössler (ed.), *Landolt–Börnstein*. New Series, Group III, vol. 17 a to i, 22 a and b, 41 A to D (Springer, Berlin, 1982–2001)
[10G1] J. Geurts, Springer Ser. Mat. Sci. **120**, 7 (2010)
[16A1] N.W. Ashcroft, N.D. Mermin, D. Wei, *Solid State Physics*, Revised edn. (Cengage Learning Asia, Singapore, 2016)

Chapter 11
Quantized Lattice Vibrations

Abstract As we have done for photons we will now attribute particle properties to the lattice vibrations. These quasi-particles, which are the normal modes of the vibrations, are called *phonons*. We will elaborate in detail the meaning of the term quasi-particles for the example of phonons and will discuss their typical properties like dispersion relation, interactions, density of states and statistics.

11.1 Phonons and the Concept of Quasi-particles

Looking again at (10.12) or (10.17) for the $u_n(t)$ we notice that these equations are rather similar to the equation of motion for an harmonic oscillator

$$m\frac{\partial^2 \xi_n}{\partial t^2} = -D\xi_n \ . \tag{11.1}$$

The only difference are the terms with indices $n \pm 1$, i.e., the off-diagonal terms in the language of matrix representation. It is now possible to find appropriate linear combinations of the u_n, so-called normal coordinates ξ_n

$$\xi_n = \sum_{n'} a_{n'} u_{n'} \tag{11.2}$$

such that the ξ_n obey equations like (11.1) or (3.5). Or, in other words, one can diagonalize the problem and introduce new normal modes (*quasi-particles*).

In doing so one usually closes the linear chain (or the three-dimensional equivalent) to a huge ring in order to close the set of differential equations. As an example, we shown in Fig. 10.9e the normal modes for the optical phonons for the Wurtzite-type structure. This crystal structure has four atoms in the basis of the primitive unit cell. These atoms (e.g. two Zn and two O for the case of ZnO) are arranged in a chain with a kink, oriented essentially parallel to the c axis. According to the rule for the number of branches of Sect. 10.3 there are nine optical modes. The patterns of motion of the various modes are indicated by the arrows. Furthermore the group-theoretical notations are given (compare Appendix B and Table B.7).

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_11

The normal modes labeled as A_1 and B_1 show elongations parallel to the **c**-axis. Consequently they are non-degenerate. The modes E_1 and E_2 exhibit elongations in the plane perpendicular to **c** and are for $\boldsymbol{k} = 0$ therefore twofold degenerate, resulting thus in a total of nine optical modes as stated above. Evidently the modes A_1 and E_1 are connected with electric dipole moments parallel and perpendicular to **c**, respectively, and can thus couple to the electric light field in the two respective polarizations.

Not all optical phonon modes can couple directly to the electromagnetic radiation field and are therefore not "optically active". They still may couple indirectly to light in Raman–scattering processes (see Sect. 12.2). Nevertheless all phonon modes above the three acoustic ones starting energetically for $\boldsymbol{k} = 0$ at a finite energy $\hbar\omega$ are called by definition *optical phonons* independent of them being *optically* or *Raman active* or both [64L1, 96G1].

A quantum mechanical treatment of the harmonic oscillator ((3.5), (11.1)) leads to discrete energy levels given by

$$E_n = \hbar\omega_n \left(n + \frac{1}{2}\right), \quad n = 0, 1, 2, 3 \ldots . \tag{11.3}$$

It is now important to note that the dispersion relation shown for example in Fig. 10.11 is not influenced by a linear transformation of the coordinates. As a consequence we may say that the lattice vibrations (more precisely the vibrations of the crystal structure) consist of quanta according to (11.3) for each wave vector $\boldsymbol{k}$ and branch i. The total energy of the lattice vibrations can be written as

$$E = \sum_{\boldsymbol{k},i} \hbar\omega_{\boldsymbol{k},i} \left(n_{\boldsymbol{k},i} + \frac{1}{2}\right) \tag{11.4}$$

in analogy to (3.8b) or (8.3), where the index i labels the various branches.

In the framework of second quantization it is possible to define creation and annihilation operators $b^{\dagger}_{\boldsymbol{k},i}$ and $b_{\boldsymbol{k},i}$ respectively, and the Hamiltonian can then again be written in terms of the number operator, in a similar way to that used already for photons

$$H = \sum_{\boldsymbol{k},i} \hbar\omega_{\boldsymbol{k},i} \left(b^{\dagger}_{\boldsymbol{k},i} b^{\dagger}_{\boldsymbol{k},i} + \frac{1}{2}\right) . \tag{11.5}$$

Obviously there is a close analogy between the quanta or particles of the electromagnetic radiation in vacuum, the photons, and the quanta of the lattice vibrations. The quanta of the lattice vibrations are called *phonons* and are considered as *quasiparticles*.

The attribute "quasi" has two reasons. In contrast to "real" particles like photons, electrons or protons, phonons can exist only in matter and not in vacuum. They are characterized by their energy and momentum $\hbar\boldsymbol{k}$. But, in contrast to vacuum, $\hbar\boldsymbol{k}$ is a *quasi-momentum* which is defined only modulo integer multiplies of the elementary translation vectors $\boldsymbol{b}_i$ of the reciprocal lattice for a crystalline solid (Sect. 7.2). The

concept of quasi-particles, which as been introduced by Landau about 50 years ago, is a very important one for the understanding of solids. In the next chapters we will become familiar with several other quasi-particles in solids. The phonons are just the first example.

The existence and the dispersion relation of phonons can be investigated by inelastic neutron scattering. An incident neutron with mass m_n from a mono-energetic beam with

$$E_i = \frac{\hbar^2 \boldsymbol{k}_i^2}{2m_n} \tag{11.6}$$

is scattered under creation or annihilation of a phonon, resulting in a neutron in the final state with E_f, $\boldsymbol{k}_f$, which are given via the conservation laws of energy and quasi-momentum

$$E_f = \frac{\hbar^2 \boldsymbol{k}_f^2}{2m_n} = E_i \pm \hbar\omega_{\text{Phonon}} \tag{11.7}$$

and

$$\boldsymbol{k}_f = \boldsymbol{k}_i \pm \boldsymbol{k}_{\text{Phonon}} + \boldsymbol{G} \tag{11.8}$$

By measuring the properties of the incident and scattered neutrons it is possible to determine the phonon dispersion relation. Note that the scattering probability for the neutrons may depend on $\boldsymbol{G}$.

As examples we show in Fig. 11.1 the dispersion relation of the phonons in two different semiconductors, Si and CdS, for various directions in $\boldsymbol{k}$-space (see Fig. 9.4). All well-known semiconductors, including the elemental ones like Si and Ge, have more than one atom per unit cell and therefore support both acoustic and optical branches. Si has only covalent binding in contrast to CdS which has a mixed ionic–covalent binding. Therefore the Si atoms do not carry an electric charge and, as a result, even the optical phonons do not couple directly to the radiation field. Consequently, at $\boldsymbol{k} = 0$ we find an oscillator strength $f = 0$ and $\Delta_{LT} = 0$ as follows from (7.26). But also in the presence of ionic binding, not all optical phonon modes couple to the electromagnetic radiation field, as discussed in connection with Fig. 10.9e.

In connection with Fig. 10.3 we have introduced an harmonic interaction potential between atoms. Actually there are strong anharmonicities, as already mentioned, which are, among others, due to the fact that the electron distribution changes almost instantaneously with the changing positions $\boldsymbol{R}_j$ of the atoms. We will use this property later for the so-called adiabatic approximation when deriving the semiconductor band structure in Sect. 14.1. This fact also results in a variation of the inter-atomic coupling (the "spring-constant" D in (11.1)) as a function of the atom spacing, i.e., in an anharmonicity.

The resulting anharmonicity manifests itself, for example, in scattering processes between phonons. We show some of them schematically in Fig. 11.2. However, one should bear in mind that all combination possibilities are exhausted only when the full three-dimensional $\boldsymbol{k}$-space is considered. In (a) a TO phonon decays under energy and momentum conservation into two acoustic phonons, while in (b) two trans-

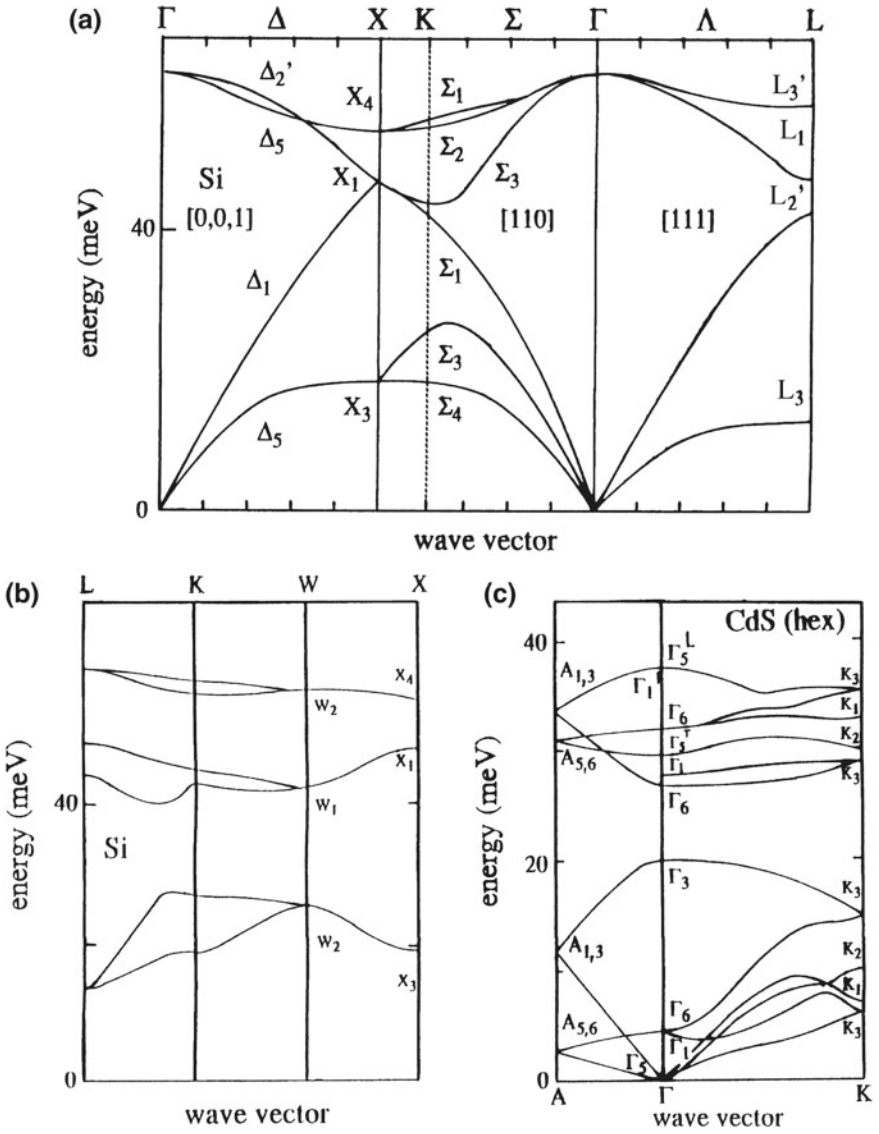


Fig. 11.1 Dispersion relation of phonons in Si and CdS for various directions in $\boldsymbol{k}$-space. Figures **a** and **c** show essentially directions which originate from the Γ point and **b** directions on the surface of the first Brillouin zone (adapted with permission from [82L1] ©1982–2001 by Springer Nature)

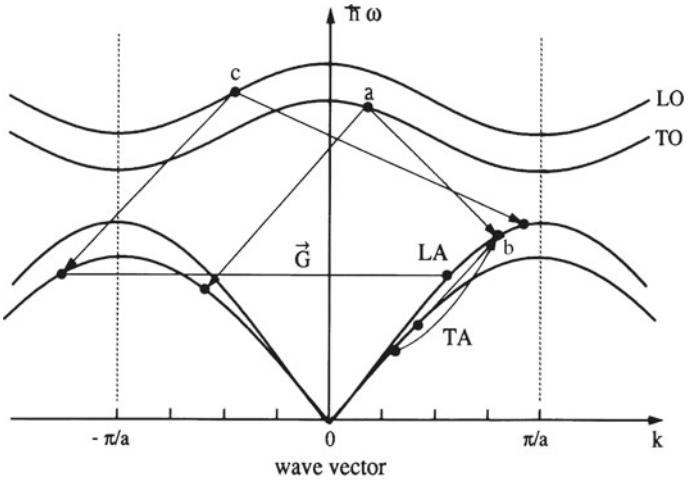


Fig. 11.2 Sketch of three possible decay or fusion processes of phonons

verse acoustic phonons combine to form a longitudinal acoustic phonon. Energy and momentum conservation read, e.g., for (a)–(c):

$$\text{(a)}\ \hbar\omega_\mathrm{i} = \hbar\omega_\mathrm{f1} + \hbar\omega_\mathrm{f2}, \quad \boldsymbol{k}_\mathrm{i} = \boldsymbol{k}_\mathrm{f1} + \boldsymbol{k}_\mathrm{f2} \tag{11.9a}$$

$$\text{(b)}\ \hbar\omega_\mathrm{i1} + \hbar\omega_\mathrm{i2} = \hbar\omega_\mathrm{f}, \quad \boldsymbol{k}_\mathrm{i1} + \boldsymbol{k}_\mathrm{i2} = \boldsymbol{k}_\mathrm{f} \tag{11.9b}$$

$$\text{(c)}\ \hbar\omega_\mathrm{i} = \hbar\omega_\mathrm{f1} + \hbar\omega_\mathrm{f2}; \qquad \boldsymbol{k}_\mathrm{i} = \boldsymbol{k}_\mathrm{f1} + \boldsymbol{k}_\mathrm{f2} + \boldsymbol{G}\ . \tag{11.9c}$$

The cases (a) and (b) are so-called n or normal processes. The decay process (c) leads into the second Brillouin zone for one phonon which is then folded back with a vector of the reciprocal lattice $\boldsymbol{G}$. If vectors of the reciprocal lattice are involved for the phonons to remain in the first Brillouin zone (e.g., in the situation (c) of Fig. 11.2) one speaks about *Umklapp-* or *u-processes*.

11.2 Density of States and Statistics of Phonons

Now we want to calculate the density of states of phonons as an intuitive example for the discussion in Chap. 4 on three-dimensional crystals. We assume for simplicity that there are two atoms per unit cell of the crystal and that the resulting three acoustic and three optical branches are degenerate, respectively, i.e., transverse and longitudinal branches coincide. Furthermore we assume an isotropic semiconductor. The resulting properties of the phonons are displayed in Fig. 11.3 where we plot on one side of the x-axis $\boldsymbol{k}$ and on the other the density of states $D(\omega)$. As elaborated in Chap. 4 the dispersion relation is needed to calculate $D(\omega)$. The linear part of the acoustic branch is described by

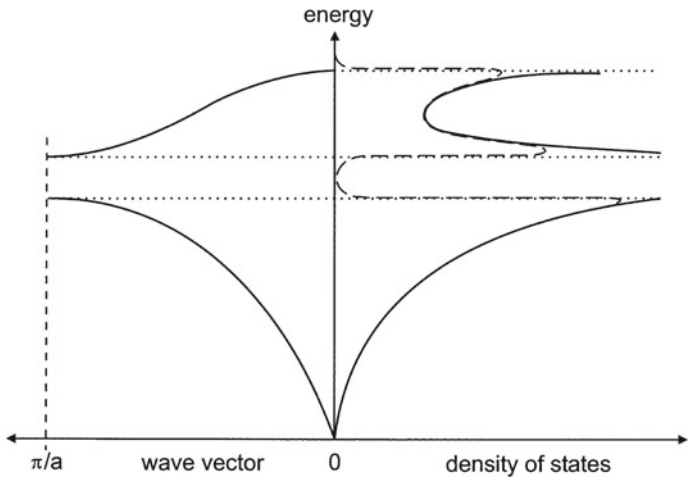


Fig. 11.3 Schematic drawing of the density of states of a three-dimensional, isotropic semiconductor with two atoms per unit cell and degenerate optical and acoustic branches

$$\omega = v_s k \ , \tag{11.10}$$

where v_s is the constant velocity of sound. This linear relation results immediately in

$$D(\omega) = \text{const} \cdot \omega^2 \tag{11.11}$$

i.e., a parabolic density of states similar to that for photons in vacuum. The difference is only in the (quite different) proportionality factors of (11.10) and (11.11).

When deviations from the linear dispersion relation start, the calculations get more lengthy and we will not go into details. From the denominator on the r.h.s. of (4.15), which is given by the derivative of the dispersion, i.e., the group velocity, we can figure out the qualitative behavior of $D(\omega)$. The density of states has in principle a singularity and in real systems a steep maximum when $\omega(\boldsymbol{k})$ has a horizontal slope as indicated in Fig. 11.3.

Note that the situation is in reality more complex than in Fig. 11.3 due to the facts, that the dispersion relation may depend on the direction of $\boldsymbol{k}$, the transverse and longitudinal branches are not degenerate, and that the first Brillouin zone is not a sphere but a polyhedron in reciprocal space. The latter has the consequence that the range of $\boldsymbol{k}$ vectors depends on the direction in the Brillouin zone. For realistic densities of states see e.g. [82L1].

In order to deduce the number of excited phonons we have to integrate over the density of states weighted with the probability that the respective states are populated. Since Phonons are bosons the commutation relation obeys

$$\left[b_{\boldsymbol{k},i}, b^\dagger_{\boldsymbol{k}',i'}\right] = b_{\boldsymbol{k},i} b^\dagger_{\boldsymbol{k}',i'} - b^\dagger_{\boldsymbol{k}',i'} b_{\boldsymbol{k},i} = \partial_{\boldsymbol{k},\boldsymbol{k}'} \partial_{i,i'} \tag{11.12}$$

where i is the label of the branch. For the statistics we have to use Bose–Einstein statistics (4.20b)

$$f_{\mathrm{BE}}(\hbar\omega, T) = \{\exp[(\hbar\omega - \mu)/k_{\mathrm{B}}T] - 1\}^{-1} . \tag{11.13}$$

Since the number of phonons is not conserved, due e.g., to phonon–phonon interaction (see above), the chemical potential μ is zero in thermal equilibrium (Chap. 4). For the total number of phonons $N(T)$ we obtain

$$N(T) = \int_0^\infty D(\hbar\omega) \left[\exp(\hbar\omega/k_{\mathrm{B}}T) - 1\right]^{-1} \mathrm{d}\omega \tag{11.14a}$$

and for the internal energy of the phonon system

$$U(T) = \int_0^\infty \hbar\omega D(\hbar\omega) \left[\exp(\hbar\omega/k_{\mathrm{B}}T) - 1\right]^{-1} \mathrm{d}\omega . \tag{11.14b}$$

Starting from (11.14b) it is easy to calculate the specific heat c_v at constant volume (which in solids is about identical to the one at constant pressure) of the phonon system

$$c_v = \frac{\partial U}{\partial T} \tag{11.15}$$

if appropriate approximations are made for $D(\omega)$. *Einstein's approximation* assumes that all phonons have the same frequency ω_{E} i.e., $D(\hbar\omega)$ is approximated by a δ-function

$$D(\hbar\omega) = \delta(\omega - \omega_{\mathrm{E}})\, 3Ns . \tag{11.16}$$

This is a good approximations in the case that optical phonons dominate the thermal properties. In *Debye's approximation* the linear part of the dispersion relation is continued up to a frequency which is also chosen to accommodate all $3Ns$ degrees of freedom of the atoms, where N is the number of unit cells and s the number of atoms per unit cell:

$$3Ns = \int_0^{\omega_{\mathrm{D}}} D(\hbar\omega)\mathrm{d}\omega \quad \text{with} \quad D(\omega) \propto \omega^2 . \tag{11.17}$$

For high temperatures (i.e., $k_{\mathrm{B}}T > \hbar\omega_{\mathrm{D}}$ or $\hbar\omega_{\mathrm{E}}$) both approximations give the classical limit, namely the *law of Dulong and Petit*

$$c_v = 3Ns\, k_{\mathrm{B}} \tag{11.18}$$

and a continuous approach to zero for $T \to 0$. In the case of low temperatures Debye's approximation better fits reality. In particular one finds the well-known T^3 law:

$$c_v \sim (T/\theta)^3 \quad \text{with} \quad k_B\theta = \hbar\omega_D \quad \text{and} \quad T < \theta \,. \tag{11.19}$$

Since the discussion of specific heat c_v is outside the scope of this book, we do not take it any further. We have introduced it simply to illustrate the applicability of the concept of the density of states. We also want to mention that the approach in (11.14b) and (11.17) is, apart from some constants and $\omega_D \to \infty$ identical to the one used to describe Planck's law of black-body radiation as seen by comparing with (4.24).

Finally, we want to point out the following. Since phonons are bosons, processes which involve phonons can be stimulated in the same way as was discussed in connection with photons in Sect. 6.1. The consequence is that in transitions involving the emission of a phonon in a certain mode, apart from other terms, a factor like

$$W_{i\to f} \sim (N^k_{Ph} + 1) \tag{11.20}$$

occurs.

In order to get a qualitative understanding we assume a lattice temperature of 77 K (i.e., the solid is cooled by liquid N_2) and an acoustic phonon state with an energy of about 0.2 meV (such phonons are typical for Brillouin scattering as will be discussed in Sect. 12.2). We find for these conditions with (11.13)

$$N^k_{Ph} \simeq 30 \gg 1 \,. \tag{11.21}$$

This means that the occupation number is much larger than one and processes which involve the emission of phonons with energies smaller than the thermal energy are stimulated by the present phonons. Depending on the process under consideration, it is however necessary to consider the reverse process, too, since it also depends on N^k_{Ph}.

Problems

11.1. Study with the data found in [82L1] the transverse and longitudinal eigenfrequencies of optical phonons of a series of semiconductors like ZnO, ZnS, ZnSe, and ZnTe, or of a similar series with the same anion and different cations. What do you conclude?

11.2. Compare the longitudinal-transverse splitting at $k = 0$ for Ge and the corresponding III–V, II–VI and I–VII compounds on the same line of the periodic table of the elements. Can you qualitatively explain the findings?

11.3. Find for Fig. 10.9e, which modes are doubly degenerate. Use your intuition together with symmetry arguments as outlined in the text and/or the tables in Appendix B.7.

References

[64L1] R. Loudon, Adv. Phys. **13**, 423 (1964). ibid. **17**, 243 (1968); A.S. Barker, R. Loudon, Rev. Mod. Phys. **44**, 18 (1972)
[82L1] O. Madelung, U. Rössler (ed.), *Landolt–Börnstein*. New Series, Group III, vol. 17 a to i, 22 a and b, 41 A to D (Springer, Berlin, 1982–2001)
[96G1] J. Geurts, Prog. Cryst. Growth Charact. **32**, 185 (1996)

Chapter 12
Phonon Polaritons

Abstract So far we have discussed the optical properties of a system of model oscillators and the basics of lattice vibrations in semiconductors. We will now demonstrate that the optical properties of phonons in semiconductors are well explained in the framework of the polariton model of Chap. 8. We will discuss the optical properties related to phonon polaritons like the phonon stop band and how to determine them via Fourier transform infrared spectroscopy. We will further describe the phonon-polartion dispersion and its experimental identification by Raman and Brillouin scattering as well as attenuated total reflection.

12.1 The Phonon Stop Band

We already mentioned that optical phonons can couple strongly to the electromagnetic field if the excitation of the optical phonon is connected with an electric dipole moment. A necessary but not sufficient precondition is, that the solid has at least partly ionic binding.

The $\boldsymbol{k}$ dependence of the phonon eigenfrequency is rather weak and covers generally only some ten meV or even less over the whole Brillouin zone. It is completely negligible if we concentrate on the region of $\boldsymbol{k}$ vectors accessible in infrared (IR) spectroscopy or scattering experiments in the visible (VIS) part of the spectrum. For IR photons in the range of the optical phonon resonances holds $k < 10^6$ cm^{-1}. Bearing in mind that the first Brillouin zone extends up to $\boldsymbol{k}$ values of around 10^8 cm^{-1} we can state, that a photon dispersion plotted into a figure of the phonon dispersion like Fig. 10.11 falls essentially onto the y-axis. We can therefore treat the *phonon polariton* according to Chaps. 7 and 8, neglecting spatial disperison. But we have to keep in mind that due to the small but finite spatial dispersion, a transverse branch extends over the whole Brillouin zone and, if appropriate (see Fig. 8.3), also a longitudinal one.

From the discussion in Chaps. 7 and 8 we know, that in an oscillator system with small damping and no spatial dispersion the propagation of polaritons is strongly suppressed between ω_T and ω_L. In this case we expect a pronounced *stop band* in the reflection spectrum similar to the calculated one shown in Fig. 7.6. These stop bands are actually found in many semiconductors.

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_12

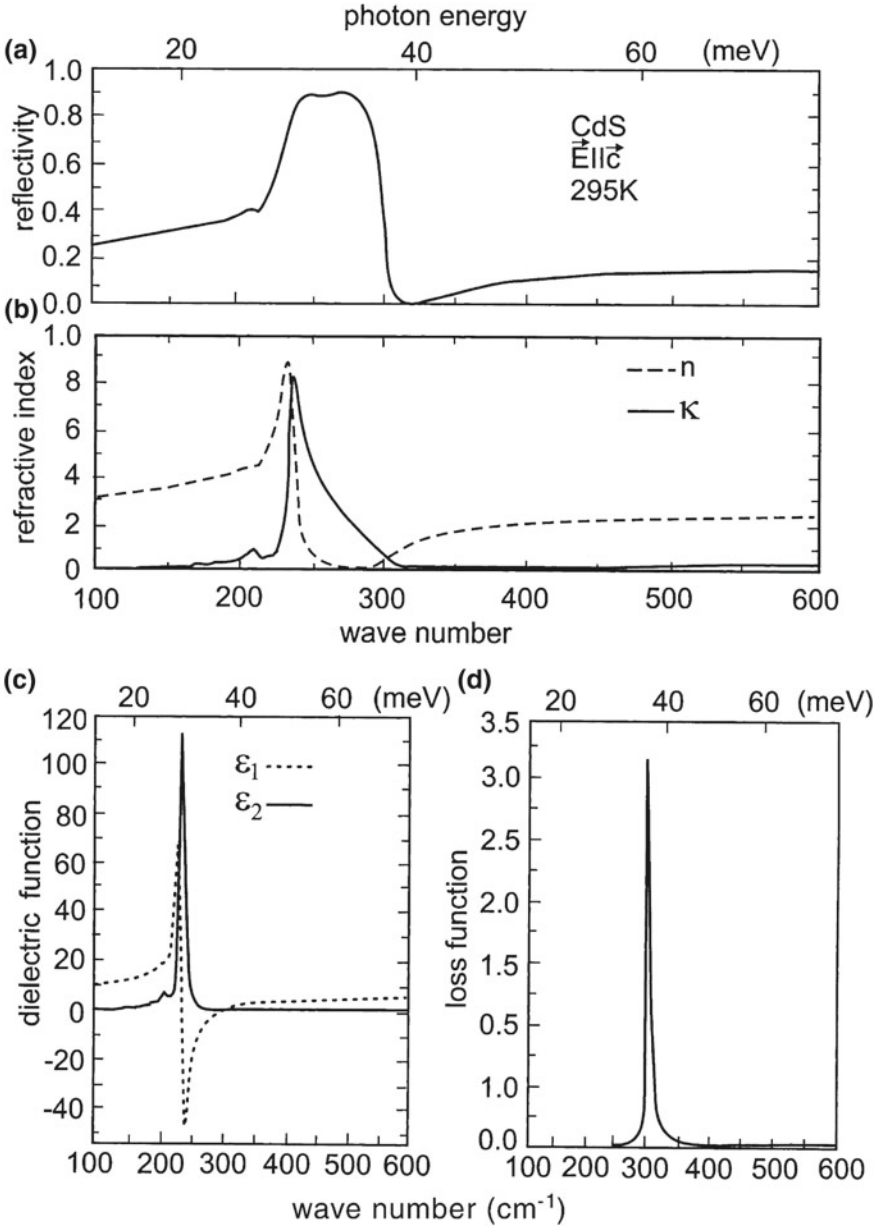


Fig. 12.1 **a** The reflection spectrum of the phonon stop band in CdS at room temperature for $\boldsymbol{E} \parallel \boldsymbol{c}$. **b** Real and imaginary parts of the complex index of refraction. **c** Real and imaginary parts of the dielectric function. **d** The loss function $\mathrm{Im}\{-1/\varepsilon(\omega)\}$ [97G1]

In Fig. 12.1a we show as an example the reflection spectrum of CdS in the IR around the phonon resonance. Spectra of other more or less ionic materials look very similar. For further examples see e.g., [81M1, 82L1, 98R1, 09K1, 10Y1]. One can clearly see the Reststrahlen band, i.e., the stop band between ω_T and ω_L as has been predicted in Fig. 7.6. The reflectivity reaches values above 0.9 at room temperature and for the polarization $\boldsymbol{E} \parallel \boldsymbol{c}$ where the Γ_1 phonons are dipole allowed (see Fig. 12.1a). The spectrum of the stop band around 250 cm^{-1} looks very similar to the one calculated using the Lorentz-oscillator model without spatial dispersion in Fig. 7.6. Please note that the small feature around 203 cm^{-1} is a *multi-phonon resonance*.

Since the reflection spectrum has been determined over a sufficiently large spectral range it is now possible with the help of the Kramers–Kronig relations and related formulae given in Appendix A to deduce the optical functions. In Fig. 12.1b we show also the real and imaginary parts of the complex index of refraction $\tilde{n}(\omega) = n(\omega) + i\kappa(\omega)$, in Fig. 12.1c the real and imaginary parts of the dielectric function $\varepsilon(\omega) = \varepsilon_1(\omega) + i\varepsilon_2(\omega)$, and in Fig. 12.1d the so-called *loss function* $\mathrm{Im}(-1/\varepsilon(\omega))$, which describes the energy loss of electrons transmitted through a thin sample. The transverse and longitudinal eigenfrequencies coincide very well with the maxima of $\varepsilon_2(\omega)$ and $\mathrm{Im}\{-1/\varepsilon(\omega)\}$, respectively.

It is obvious, that the experimental and deduced spectra coincide nicely with the results of the model calculations displayed in Figs. 7.3 and 7.5. Transmission spectra of thin samples also agree with Fig. 12.1 as shown, e.g., in [97G1]. This agreement allows us to claim good understanding of the optical properties of phonons within the Lorentz-oscillator model.

Experimental Technique: Fourier Transform Infrared Spectroscopy (FTIR)

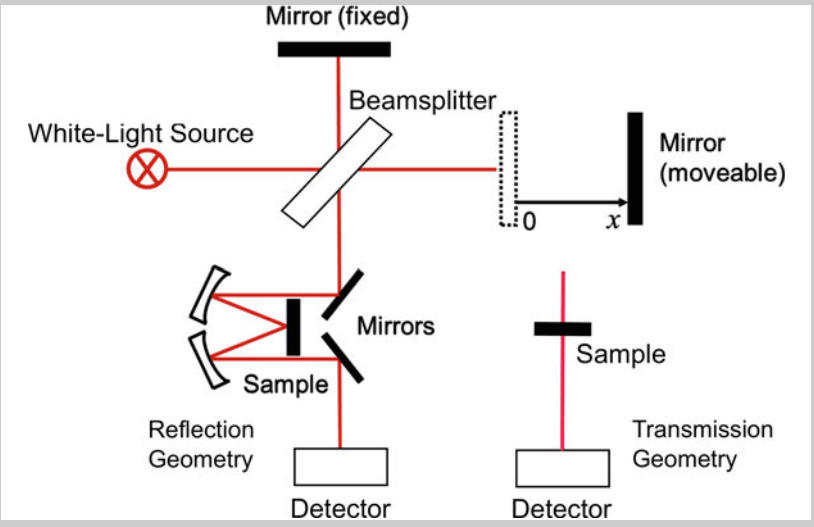


Fig. FTIR-1 Setup of a FTIR experiment for absorption and reflection measurements

The optical spectra (reflectivity, transmission) of phonons which are spectrally situated in the far infrared are typically recorded using *Fourier Transform Infrared Spectroscopy (FTIR)*. This is an interferometric technique which is used not only since suitable refractive and dispersive optical elements for spectroscopy are not available. There are actually some advantages of this method with respect to classical spectroscopy.

The setup for this method (Fig. FTIR-1) is based on a Michelson interferometer with one movable mirror. The light source is a broadband emitter (typically black-body radiation, e.g., from a glowbar or a halogen lamp). The sample is placed outside the interferometer before the detector in absorption or reflection geometry. The detector records the interferogram of the two co-linear light beams from the two arms after being transmitted through or reflected from the sample.

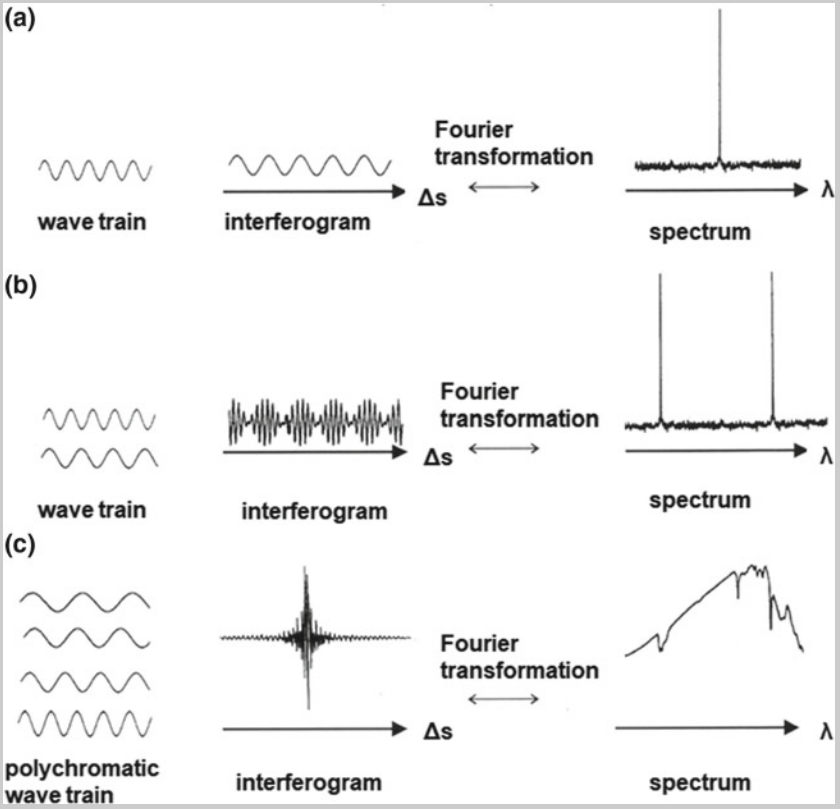


Fig. FTIR-2 Interferograms and their Fourier transformation into a spectra $I(\lambda)$ for a light source emitting at a single wavelength (**a**), at two wavelengths (**b**) and emitting a continuous spectrum (**c**)

The total intensity at the detector as a function of optical pathlength difference Δs (seen in the second column of Fig. FTIR-2) represents essentially the auto-correlation of the light impinging from the sample. The intensity oscillates cosine-like around a value given by half of the incident intensity. This is the signal for large negative and positive pathlength differences Δs where this difference is larger than the coherence length of the light. Coherence is maximum at $\Delta s = 0$ (called the white-light point) and the recorded intensity has its highest peak. The spectrum $I(\lambda)$ is calculated by Fourier transformation of $I(\Delta s)$. Interferograms and transformed spectra are shown for typical situations in Fig. FTIR-2.

With the FTIR method one can measure spectra over large spectral ranges with very high resolution. This also works when the spectrum has segments of law intensity since during the whole measurement time, the complete intensity falls onto the detector and is evaluated (called throughput or energy advantage). Further the complete spectral range is measured simultaneously, not wavelength by wavelength (multiplex advantage).

Typical applications of FTIR in solid-state spectroscopy are the determination of phonon parameters from their stop bands (see Fig. 12.2 for examples), to characterize superlattices from folded acoustic phonon or confined optical phonon resonances, the determination of the plasma edge in doped semiconductors or spectroscopy of plasmonic excitations. For reviews of the method and its applications see, e.g., [72B1, 91Z1, 98G1, 06A1, 09K1].

Concerning optical spectroscopy of phonons we should recall some details of the optical properties of phonons. Optical phonons in semiconductors with purely covalent binding like Si or Ge have, at $k = 0$, zero longitudinal-transverse splitting and zero oscillator strength, as was shown in Fig. 11.1. Furthermore, in samples with (partly) ionic binding, optical phonon modes may exist, which do not carry a dipole moment. In both cases the optical phonons do not couple to the electromagnetic field, i.e., they cannot be investigated by optical spectroscopy in the spectral region of the phonon resonances.

Phonons that do not couple directly to light are said to be not *optically active*. For ZnO these are the modes B_1 and E_2 (Γ_3 and Γ_6) in Fig. 10.9e, while the A_1 and E_1 (Γ_1 and Γ_5) modes carry a dipole moment parallel and perpendicular to the crystallographic $\boldsymbol{c}$-axis, respectively. Optical phonons that are not optically active might still be *Raman active*, i.e., they lead to Raman lines in Raman spectroscopy (see 12.2).

In order to achieve in the stop band a reflectivity near unity the imaginary part $\kappa(\omega)$ has to dominate in (7.35a). This dominance can be suppressed when thin layers of a semiconductor sample are investigated. In particular, when the sample thickness is small compared to the penetration depth of the evanescent light field in the stop-band region calculated from the inverse of κ, then the light can change from an evanescent to an oscillatory propagating mode after passage of the thin layer. The resulting modifications of the stop-band spectra are shown in Fig. 12.2 [06A1].

For a thin layer of ZnSe with a thickness of about 200 nm the stop band has changed from the shape of Fig. 7.6 to a narrow line at $\hbar\omega_0$. This line coincides with the peak in $\kappa(\omega)$ shown in the insert if Fig. 12.2b. For frequencies above this peak the penetration depth strongly increases due to the decreasing κ and the light can penetrate into the (at this frequency range transparent) GaAs substrate. The stop band of the thick GaAs substrate is well developed. The phonon features of both materials show with increasing temperature a broadening due to increasing damping and a small red shift due to increasing lattice constants.

The consequences of a transition to a thicker ZnSe layer are shown in Fig. 12.2b. The ZnSe stop band broadens (upper row right) due to a decreasing penetration depth. The spectra are well explained by fits based on the Lorentz-oscillator model. In the lower row of Fig. 12.2b the same model was used to simulate the reflection spectra now of a very thick (1,000 μm) layer of ZnSe on top of a GaAs substrate. For finite damping only the stop band of ZnSe is seen now, because the high-frequency absorption tail in the ZnSe layer prevents the light from reaching the GaAs substrate. The situation becomes different for an assumed vanishing damping. The reflection features obtain the "rectangular" shape of Fig. 7.6 for this case and spectrally close lying Fabry-Perot modes appear due to the finite thicknesses of layer and substrate. Further the reflection feature of the GaAs is still seen, because now the imaginary part of the refractive index is zero below the transverse and above the longitudinal eigenfrequencies.

12.2 The Dispersion of Phonon Polaritons

The phonon stop band gives indirect evidence of the polariton properties of optical phonons. But, the polariton dispersion can also be addressed directly by Raman scattering. Before we review this type of scattering experiments, which require strong-coupling conditions, we want to discuss Raman scattering in the more common version which can be described in the weak-coupling limit.

First, it is important to understand the different role of phonons in IR and Raman spectroscopy, respectively. Infrared-active Phonons directly couple via their electric dipole moment to the incident light. This results in the formation of phonon polaritons and the resonant behavior of the dielectric function (Fig. 7.3) at far-infrared frequencies. Raman-active phonons modulate via their (periodic) local density fluctuations the polarizability of the crystal. Thus the dielectric function at the incident light frequency (typically in the visible or near- infrared) is modulated with the phonon frequency resulting in side bands of the incident light wave shifted up or down by this phonon frequency [09K1, 10Y1]. From symmetry arguments follows that in crystals with inversion symmetry, optical phonons are either Raman or IR active, i.e., the selection rules complement each other. In samples without inversion symmetry, optical phonons may be both Raman and IR active.

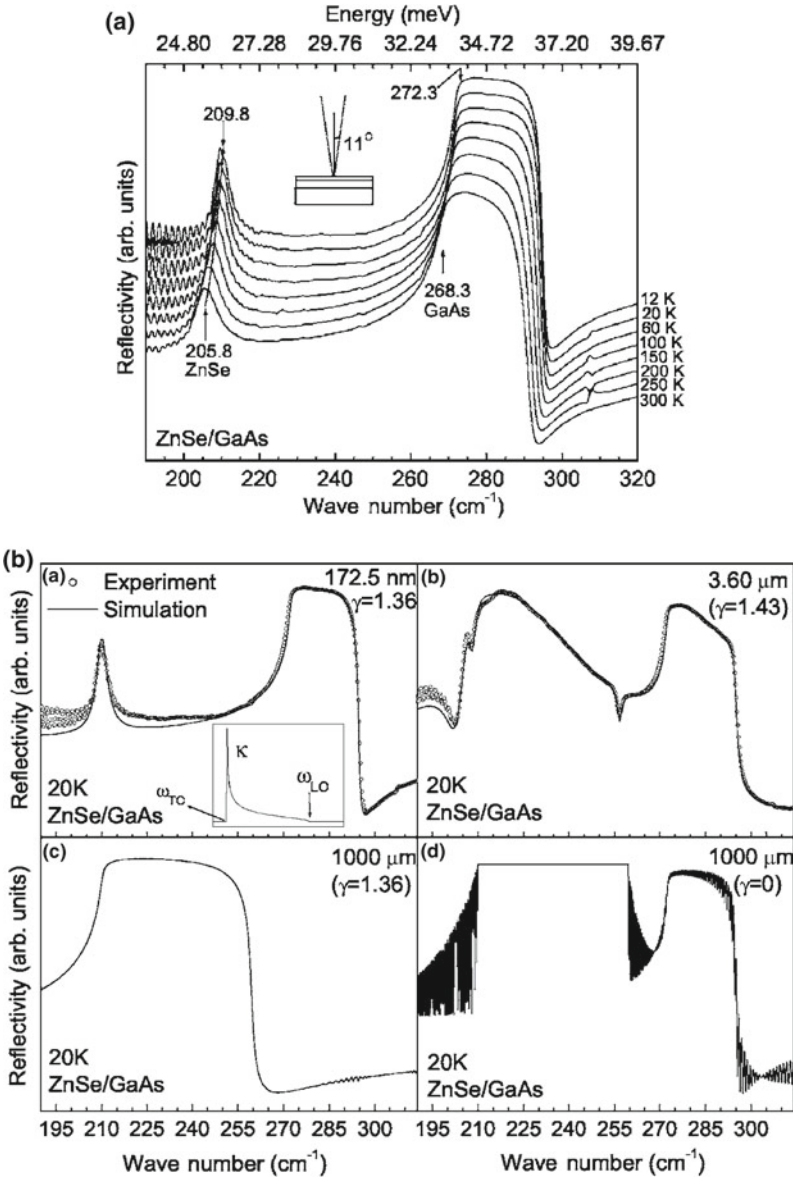


Fig. 12.2 Measured reflection spectra of a GaAs substrate covered with a ZnSe layer of 172.5 nm thickness for various temperatures (**a**). The 20 K spectrum of this layer and a layer with 3.6 μm thickness (**b**) together with theoretical fits with the indicated values of the damping γ (*upper row*) together with the imaginary part in the insert and calculated spectra for a thickness of the ZnSe layer of 1,000 μm with $\gamma = 1.36$ (**c**) and for vanishing damping (**d**) (reprinted with permission from [06A1] ©2006 American Physical Society https://doi.org/10.1103/PhysRevB.73.045211)

Experimental Techniques: Raman and Brillouin Scattering

Raman scattering is the inelastic scattering of light under emission or absorption of an optical phonon or a phonon polariton for the case, that the phonon mode under consideration is both Raman and IR active. One can also say that a photon creates virtually an electronic excitation (exciton or an electron–hole band-to-band transition) and is scattered by emission or absorption of an optical phonon. But this model fails in some respects as we will see below.

In a Raman-scattering process energy and (quasi-)momentum have to be conserved, i.e.,

$$\hbar\omega_{\mathrm{R}} = \hbar\omega_{\mathrm{i}} \pm \hbar\Omega_{\mathrm{Ph}} \quad \text{and} \quad \boldsymbol{k}_{\mathrm{R}} = \boldsymbol{k}_{\mathrm{i}} \pm \boldsymbol{k}_{\mathrm{Ph}}\,, \tag{RS-1}$$

where the index i stands for the initial or incident light quantum, R for the Raman signal and Ph for the created or annihilated phonon. The creation of a phonon, represented by the – sign, is called *Stokes scattering*, while the annihilation of a phonon, represented by the + sign, is termed *anti-Stokes scattering*.

Usually one chooses ω_{i} and ω_{R} in the transparent spectral region of the semiconductor well above the phonon resonance and below the band-to-band transition. Often $\hbar\omega_{\mathrm{i}}$ is determined by the emission lines of a readily available laser such as an Ar^{+} laser.

Equation (RS-1) results in values around 10^5 cm^{-1} for $\boldsymbol{k}_{\mathrm{i}}$ and $\boldsymbol{k}_{\mathrm{R}}$. In a backward or 90° scattering geometry $\boldsymbol{k}_{\mathrm{Ph}}$ is also around 10^5 cm^{-1} and thus clearly lies on the phonon-like part of the dispersion relation (see, e.g., Fig. 12.4) or on the longitudinal branch. In Fig. RS-1a we plot the dispersion of acoustic phonons, of transverse optical phonon polaritons and of a longitudinal optical phonon starting from the position of the incident photon $\hbar\omega_{\mathrm{i}}$; $\boldsymbol{k}_{\mathrm{i}}$ to the left (backward scattering) and into the negative energy direction (Stokes scattering). Actually the incident photon is also a polariton since it couples to the polarization density of electronic resonances. This is illustrated here by the fact that the linear photon-like polariton dispersion in Fig. RS-1a starts to bend over to an exciton-like dispersion close to an exciton resonance. In Fig. RS-1b we also give a typical Raman spectrum.

This case is covered by the weak-coupling picture. The selection rules are given in a classical description by the Raman tensor (see, e.g., [09K1, 10Y1]). It couples the polarization directions of the incident and scattered photons and yields the scattered intensity from the modulation of the suszeptibility induced by the phonon amplitude. A comprehensive treatment including phonon generation requires quantum mechanical evaluation of transition matrix elements [10Y1]. The group theory presented in Appendix B gives in this case information about which transitions are allowed and which are not.

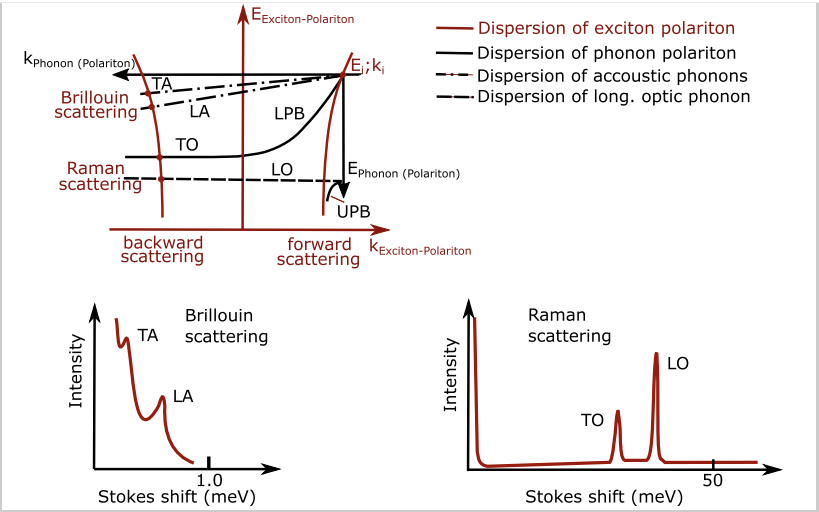


Fig. RS-1 Schematic drawing of energy and momentum conservation in (RS-1) for a Raman scattering process (Stokes emission) in a backscattering configuration (adapted with permission from [71M1] ©1971 American Physical Society https://dx.doi.org/10.1103/PhysRevLett.26.86 (**a**), and typical spectra for Brillouin (**b**) and Raman scattering (**c**). Note the different scales on the x-axes in (**b**) and (**c**)

An easily intelligible formula can be derived for the intensity ratio of Stokes to anti-Stokes emission, i.e., scattering under emission or absorption of a phonon and assuming that the density of final states for Stokes and anti-Stokes emissions are practically identical:

$$\frac{I_\mathrm{s}}{I_\mathrm{a}} = \left(\frac{\omega_\mathrm{i} - \Omega_\mathrm{Ph}}{\omega_\mathrm{i} + \Omega_\mathrm{Ph}}\right)^4 \cdot \frac{N_\mathrm{Ph} + 1}{N_\mathrm{Ph}} = \left(\frac{\omega_\mathrm{i} - \Omega_\mathrm{Ph}}{\omega_\mathrm{i} + \Omega_\mathrm{Ph}}\right)^4 \cdot \exp\left\{\frac{\hbar\Omega_\mathrm{Ph}}{k_\mathrm{B}T}\right\}, \quad \text{(RS-2)}$$

This ratio just reflects the fact that the phonon emission is proportional to $(N_\mathrm{ph} + 1)$ and the absorption proportional to N_Ph, where N_Ph is the phonon occupation number of the respective mode.

The first term incorporates the scattering probability proportional to ω^4. This factor is usually close to unity because $\omega_\mathrm{i} \gg \Omega_\mathrm{Ph}$.

Brillouin scattering is the analog of Raman scattering for the case of acoustic phonons (for details see [09K1, 10Y1]). Because of the rather flat dispersion relation of acoustic phonons, with a slope given by the velocity of sound (instead of c/n), one finds even in a back-scattering configuration only much smaller (anti-) Stokes shifts, which are usually $\leq$1 meV (Fig. RS-1c). Since the coupling of acoustic phonons to photons is also much weaker than for optical phonons, some high-resolution techniques and an efficient suppression of stray light are necessary to detect Brillouin scattering, as can be seen by comparing the abscissas of Fig. RS-1b and c. Since the dispersion relation of acoustic phonons is linear in the range of interest, (RS-1) can be rewritten for

Brillouin scattering in a back-scattering configuration as

$$\frac{\Omega_{\mathrm{Ph}}^{\mathrm{LA,TA}}}{v_{\mathrm{s}}^{\mathrm{LA,TA}}} = \frac{\omega_{\mathrm{i}} - \omega_{\mathrm{R}}^{\mathrm{LA,TA}}}{v_{\mathrm{s}}^{\mathrm{LA,TA}}} = 2\frac{\omega_{\mathrm{i}}}{c} n(\omega_{\mathrm{i}}) \quad . \tag{BS-1}$$

Here we assume that the incident light is so far away from any electronic resonance that the real part of the refractive index $n(\omega)$ does not change significantly over the Brillouin shift. In this case either $n(\omega)$ or v_{s} can be determined from Brillouin scattering if the other quantity is known. If ω_{i} approaches the resonance of the exciton polariton, Brillouin scattering is an efficient means of $\boldsymbol{k}$-space spectroscopy of these resonances. We will come back to this aspect in Sect. 21.1.4.

To get access to the phonon-polariton dispersion and the strong-coupling case the scattering geometry of the Raman experiment is changed from a backward (or 90°) geometry to forward scattering. As a consequence $\boldsymbol{k}_{\mathrm{Ph}}$ can become small and can fall for IR and Raman active modes into the transition region from phonon-like to photon-like polaritons. Experiments in this geometry thus allow measuring the dispersion of the phonon polariton rather directly.

Experiments of this type (see e.g., [65H1, 72B1, 75C1, 80C1], are performed by collecting the scattered light on the entrance slit of a spectrometer such that the height on the slit is a measure of $\boldsymbol{k}_{\mathrm{i}} - \boldsymbol{k}_{\mathrm{R}}$ (Fig. 12.3a). The wavelength dispersion of the spectrometer then gives the Ω axis. We show an example schematically in Fig. 12.3b.

Figure 12.4 gives the dispersion relation for the phonon polariton in GaP reconstructed from this type of experiments, which of course, has to take into account the refraction of the beams at the surface of the sample in contrast to the simplifications in Fig. 12.3a. These experiments can obviously only be understood in the strong-coupling or polariton picture. In [72B1, 09D1] it has been addressed explicitly that one can follow the phonon-polariton dispersion well beyond the turning point, obtained for a simple Lorentz oscillator with finite damping (compare Fig. 8.1a).

In another group of experiments, the dispersion of phonon polaritons has been investigated or verified by an inelastic scattering of two light quanta in which the differences of the wave vectors and of the quantum energies match a point on the dispersion curve of the phonon polariton [98B1, 99R1, 01W1, 02C1].

Also Raman processes are possible in which two phonons are created (or annihilated), e.g., due to the strong anharmonicity of the potential between atoms resulting in phonon–phonon interaction mentioned in connection with (11.9a)–(11.9c). If this is a two-step process, in the sense that the Raman polariton ω_{R} in (RS-1) undergoes a second Raman scattering process. Then both phonons have small wavevectors compared to the size of the Brillouin zone. If, on the other hand, both phonons are emitted simultaneously, only the sum of the phonon wavevectors has to fulfill (RS-1). The individual phonons can come from any part of the Brillouin zone. Consequently the

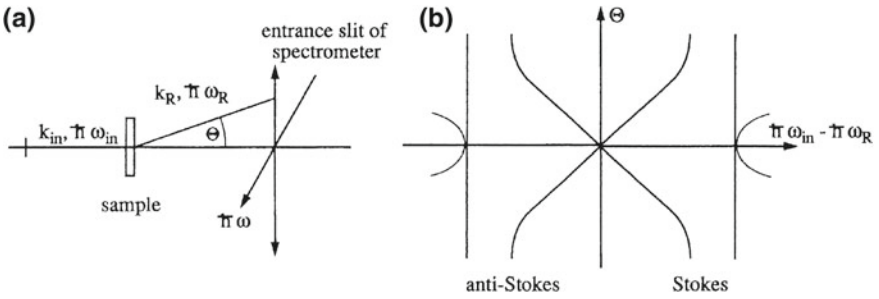


Fig. 12.3 Schematics of a Raman scattering experiment in the forward direction with a spectrometer (**a**) and the visualization of the relation between Θ and the phonon energy $\hbar\Omega$ in the output plane of the spectrometer (**b**)

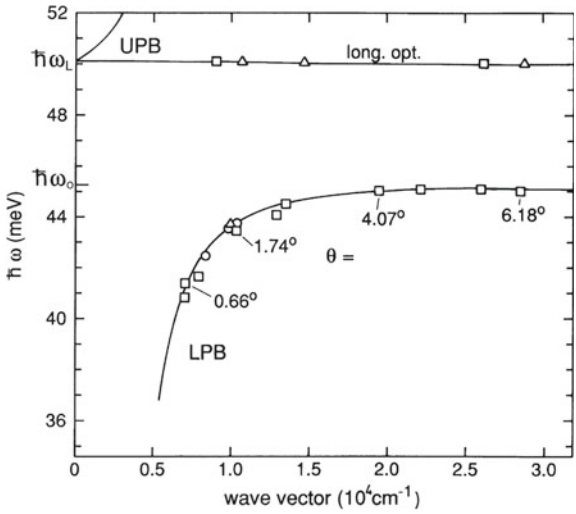


Fig. 12.4 The dispersion of the phonon polariton in GaP, measured by the method shown in Fig. 12.3 reprinted with permission from [65H1] ©1965 American Physical Society https://dx.doi.org/10.1103/PhysRevLett.34.1516

Raman spectrum then reflects to a certain extent the density of states of the phonons. Details of such processes are given in [79S1, 94C1, 03H1].

We want to conclude this section with some comments on the dynamics of phonons since the the discussion in the second volume of this book will focus on the dynamics of electronic excitations only.

Coherent Phonon polaritons can be excited in coherent Raman scattering experiments using picosecond laser pulses or by impulsive excitation using (femto-second) laser pulses which are shorter than the oscillation period of the phonon (for reviews see, e.g., [92K1, 97M1, 08F1, 03W1]. There are different excitation mechanisms possible like impulsive stimulated Raman scattering (ISRS) in Ge or screening of surface fields by generation of electron–hole pairs in polar semiconductors like GaAs. These phonons lead to Raman scattering in the time domain which can be detected by ultra-fast oscillations in the reflectivity or transmission in the visible [08F1]. They

also can be investigated by their emission of electromagentic radiation in the THz regime [92K1, 98T1, 03W1]. Phonon polaritons with well defined $\boldsymbol{k}$ can be generated by interference of two crossing pump beams or by optical rectification/difference-frequency generation. The generation of the THz emission is described in terms of Cherenkov radiation [03W1].

The dephasing time of the coherent phonons, which is related to the homogeneous linewidth of the phonon resonance via an uncertainty relation, can be deduced for IR-active optical phonons from a fit of the reflection spectra as in Figs. 12.1 and 12.2 assuming that only homogeneous broadening matters [06A1]. A direct measure of the dephasing time is given by the temporal decay of the above mentioned oscillations, which is typically in the picosecond regime. Effects which contribute to the homogeneous broadening are the anharmonicity of the potentials between (neighboring) atoms, degeneracy of the densities of state of one- and two-phonon states [03S1] and interaction with free carriers [92K1].

Also coherent longitudinal acoustic phonons can be studied by optical experiments or are used in optical experiments to investigate the structural properties of the sample. These phonons are created by ultra-short laser pulses at the sample surface and propagate into the depth. Interference of reflected probe-pulse light from the surface with light reflected from the plane wave fronts of the propagating phonon leads to long-lasting (several hundreds of picoseconds) oscillations in the reflectivity (see, e.g., [86T1, 08X1, 15R1] for further reference).

We can state that coherent phonons are versatile tools to study e.g., structural properties of semiconductors, mechanisms of phonon-assisted relaxation of carriers in low-dimensional semiconductor structures, interaction of carriers like an non-equilibrium electron–hole plasmas with phonons or coupled plasmon-phonon modes in doped semiconductors [92K1, 97M1, 08F1, 10G1].

12.3 Surface Phonon Polaritons

We already discussed in Sect. 8.6 that surface polariton modes exist in the range between the transverse and longitudinal eigenmodes. These modes can be observed by *attenuated total reflection (ATR)*, i.e., by coupling the evanescent wave shown in Fig. 5.3 to the material under investigation. More details of this technique are given in Sect. 21.1.5.

Here we show in Fig. 12.5a ATR spectra and in b the deduced dispersion relation of the *surface phonon polariton* in GaP. The agreement with the schematic drawing of Fig. 8.14a is obvious. For more details see [74B1, 97V1].

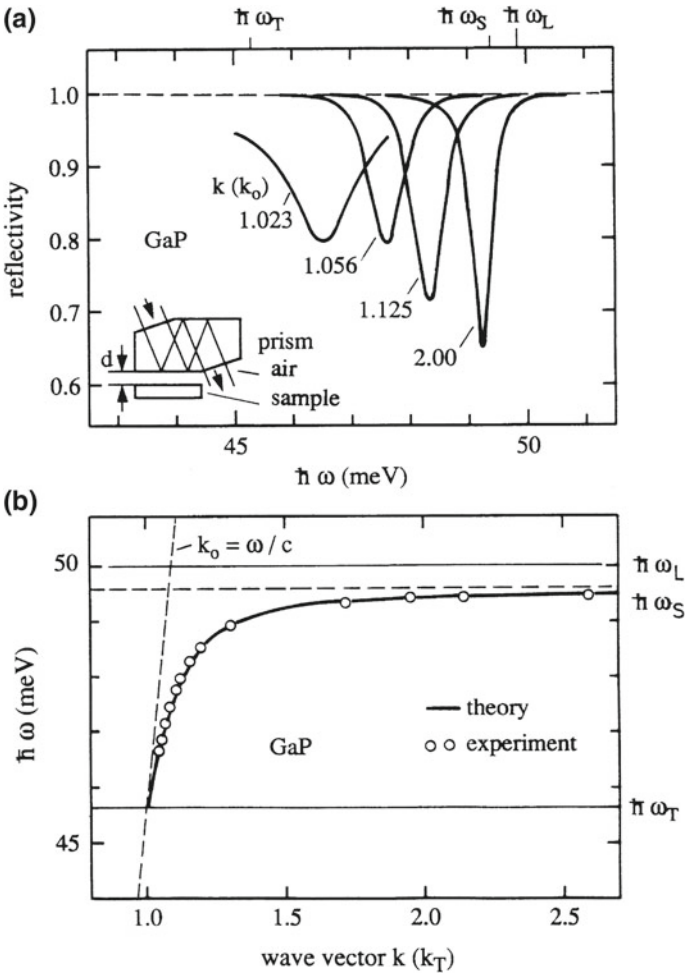


Fig. 12.5 Spectra of attenuated total reflection (**a**) in the region of the optical phonon resonances of GaP and the resulting dispersion relation of the surface phonon polariton (**b**). Compare with Fig. 12.4 for the bulk modes (adapted with permission from [72M1] ©1972 American Physical Society https://dx.doi.org/10.1103/PhysRevLett.28.811)

Problems

12.1 Try to find more reflection spectra like that in Fig. 12.1b in the literature. Deduce $\hbar\omega_T$, Δ_{LT}, ε_s and ε_b from these spectra and compare with values in the literature.

12.2 Show that the (eventually only weak) dependence of ω on $\boldsymbol{k}$ is important to explain the experimental fact that the TO and LO phonon modes can be followed

through the whole Brillouin zone by neutron scattering. Compare for the explanation with Fig. 8.1 for vanishing and finite damping.

12.3 Why is the phonon spectrum of high T_c superconductors so complex?

12.4 Which trend would you expect for the zone boundary LA and TA phonons when going from ZnO via ZnS and ZnSe to ZnTe. Compare with data in the literature. What do you expect for zone center optical phonons?

12.5 Can you give qualitative arguments for why the optical phonons in $CdS_{1-x}Se_x$ and $Zn_{1-y}Cd_yS$ are of the persistent and amalgamation type, respectively? Consider the atomic masses of the oscillating atoms.

12.6 Find out why phonon data for CdS and CdSe obtained by neutron scattering are relatively rare?

References

[65H1] C.H. Henry, J.J. Hopfield, Phys. Rev. Lett. **15**, 964 (1965)
[71M1] R.M. Martin, T.C. Damen, Phys. Rev. Lett. **26**, 86 (1971)
[72B1] A.S. Barber Jr., R. Loudon, Phys. Rev. **44**, 18 (1972)
[72B1] J. Bell, *Introductory Fourier Transform Spectroscopy* (Academic, New York, 1972)
[72M1] N. Marschall, B. Fischer, Phys. Rev. Lett. **28**, 811 (1972)
[74B1] G. Borstl, H.J. Falge, A. Otto, *Surface and Bulk Phonon Modes Observed by Attenuated Total Reflection*. Springer Tracts in Modern Physics, vol. 74 (Springer, Berlin, 1974)
[75C1] R. Claus, L. Merten, J. Brandmüller, *Light Scattering by Phonon-Polaritons*. Springer Tracts Modern Physics, vol. 75 (Springer, Berlin, 1975)
[79S1] R.L. Schmidt et al., Phys. Rev. B **20**, 3345 (1979)
[80C1] R. Claus, Phys. Status Solidi (b) **100**, 9 (1980)
[81M1] O. Madelung, *Introduction to Solid State Theory*. Springer Series in Solid State Sciences, vol. 2 (Springer, Berlin, 1981)
[82L1] O. Madelung, U. Rössler (ed.), *Landolt–Börnstein*. New Series, Group III, vol. 17 a to i, 22 a and b, 41 A to D (Springer, Berlin, 1982–2001)
[86T1] C. Thomsen, H.T. Grahn, H.J. Maris, J. Tauc, Phys. Rev. B **34**, 4129 (1986)
[91Z1] B. Zhang et al., Semicond. Sci. Technol. **6**, 822 (1991)
[92K1] W. Kütt, Festkörperprobleme/Adv. Solid State Phys. **32**, 133 (1992)
[92K1] W.A. Kütt, W. Albrecht, H. Kurz, IEEE J. Quantum Electron. **28**, 2434 (1992)
[94C1] M.P. Chamberlain, C. Trallero-Giner, M. Cardona, Phys. Rev. B **50**, 1611 (1994)
[97G1] M. Göppert, Diplom thesis, Universität Karlsruhe (TH) (1997)
[97M1] R. Merlin, Solid State Commun. **102**, 207 (1997)
[97V1] R.V. Velasco, F. Garcia-Moliner, Surf. Sci. Rep. **28**, 123 (1997)
[98B1] H.J. Bakker, S. Hunsche, H. Kurz, Rev. Mod. Phys. **70**, 523 (1998)
[98G1] M. Göppert et al., Phys. Rev. B **57**, 13068 (1998)
[98R1] T. Ruf, *Phonon Raman Scattering in Semiconductors Quantum Wells and Superlattices*. Springer Tracts in Modern Physics, vol. 142 (Springer, Berlin, 1998)
[98T1] M. Tani et al., J. Appl. Phys. **83**, 2473 (1998)
[99R1] V. Romero-Rochin et al., J. Chem. Phys. **111**, 3559 (1999)
[01W1] T. Watanuki et al., J. Phys. Soc. Jpn. **70**, 2784 (2001)
[02C1] T.F. Crimmins, N.S. Stoyanov, K.A. Nelson, J. Chem. Phys. **117**, 2882 (2002)

[03H1] U. Haboeck et al., Phys. Status Solidi C **0** (6), 1710 (2003)
[03S1] J. Serrano et al., Phys. Status Solidi B **235**, 260 (2003)
[03W1] J.K. Wahlstrand, R. Merlin, Phys. Rev. B **68**, 13 (2003)
[06A1] K.C. Agarwal et al., Phys. Rev. B **73**, 045211 (2006)
[08F1] M. Först , T. Dekorsy, in *Coherent Vibrational Spectroscopy*, ed. by S. de Silvestri, G. Cerullo, G. Lanzani (2008), p. 129
[08X1] Y. Xu et al., Phys. Status Solidi C **5**, 2632 (2008)
[09D1] V.Yu. Davidov et al., Phys. Rev. B **80**, 081204(R) (2009)
[09K1] H. Kuzmany, *Solid State Spectroscopy*, 2nd edn. (Springer, Berlin, 2009)
[10G1] J. Geurts, in *Zinc Oxide: From Fundamental Properties Towards Novel Applications*. Springer Series in Materials Science, vol. 120 (Springer, Heidelberg, 2010), p. 7
[10Y1] P.Y. Yu, M. Cardona, *Fundamentals of Semiconductors*, 4th edn. (Springer, Heidelberg, 2010)
[15R1] P. Ruello, V.E. Gusev, Ultrasonics **56**, 21 (2015)

Chapter 13
Phonons in Solids of Perturbed Lattice Periodicity

Abstract In this chapter we treat the properties of phonons in scenarios where the periodicity of the crystalline lattice is perturbed. This is the case for crystals with a random distribution of atoms on lattice sites (alloys) or for local deviations from periodicity (defects). We also consider the situation of reduced periodicity in layered structures like superlattices and quantum wells. Finally we address confinement of phonons in quantum dots.

13.1 Phonons in Alloy Semiconductors

An *alloy* is a random binary, ternary or higher mixture of atoms. In contrast to amorphous materials, the atoms in an alloy are still sitting on well defined, periodically arranged lattice sites. Many semiconductors form ternary (like $Al_{1-y}Ga_yAs$) or quaternary alloys (like $Zn_{1-y}Mg_yS_{1-x}Se_x$) (for a review see [09A1]). In some cases the alloys exist for the whole composition range $0 \leq x,\ y \leq 1$. In other cases, there is a miscibility gap.

Two widely investigated semiconductor alloys are $Al_{1-y}Ga_yAs$ and $CdS_{1-x}Se_x$. The first one crystallizes in the cubic zinc-blende structure the second one usually in the hexagonal wurtzite type. However, the two different cations Al and Ga or the two anions S and Se are randomly distributed over the respective sub-lattice with an average concentration y or x. The properties of such alloys are often treated by a so-called *virtual crystal approach* (see also Sect. 18.2 for electronic properties of alloy semiconductors). Here a periodic lattice is constructed from unit cells of atoms with averaged properties according to the concentration y or x.

What are the consequences of the distorted lattice periodicity in alloys for the phonon properties? For long-wavelength acoustic phonons the situation is generally rather simple. The dispersion relation starts linear as in Figs. 10.5, 10.7, 10.10 or Fig. 10.11. The slope or velocity of sound is a weighted average of the ones of two parent components (GaAs and AlAs or CdS and CdSe, respectively). Long wavelength means large compared to the inter-atomic distance. When approaching the boarder of the first Brillouin zone, things become less clear. Since only one sub-lattice is elongated (see Fig. 10.9b) the phonon resonance may be broadened

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_13

depending on which type of atom oscillates due to their different masses and coupling constants to their neighbors.

In this sense even a chemically pure crystal, containing however different isotopes of the same element, has to be considered as an alloy with a weak disordered. The resulting broadening of the phonon resonances can be eliminated in isotope-pure materials. For details of this topic see e.g., [94C1, 94R1].

For the optical phonons, two different situations may arise. The reason is essentially given by the physics of coupling between oscillators with approximately equal or clearly different resonance frequencies. The latter occurs if the phonon bands of the parent compounds in an alloy $AB_{1-x}C_x$ do not overlap energetically, i.e., the AB phonon modes do not find a resonant partner of AC modes for any $\boldsymbol{k}$ vector and vice versa. Such alloys exhibit generally the so-called *persistent mode* behavior. This means, that the eigenfrequencies of the pure AB and AC modes vary only marginally under alloying as a function of x, but their relative weights and oscillator strengths change. This is shown for the case of TO and LO phonons in $CdS_{1-x}Se_x$ in Fig. 13.1a. One sees clearly that the frequencies change only slightly, but the longitudinal-transversal splitting Δ_{LT} which is proportional to the oscillator strength varies continuously and complementary with composition.

If the bands of e.g., the LO- phonons overlap energetically in the two parent materials as is the case e.g., in $Zn_{1-y}Cd_yS$ one finds the so called *amalgamation type* behavior. Here the LO eigenfrequency shifts continuously with the composition y from the one of AB to the one of AC in an alloy $AB_{1-x}C_x$ as shown in Fig. 13.1b.

In addition to the limiting cases of amalgamation and persistent mode type, there is also an intermediate mode behavior. Furthermore, a weak mode may result from a local weakening of bonds. For these cases see e.g., [06A1].

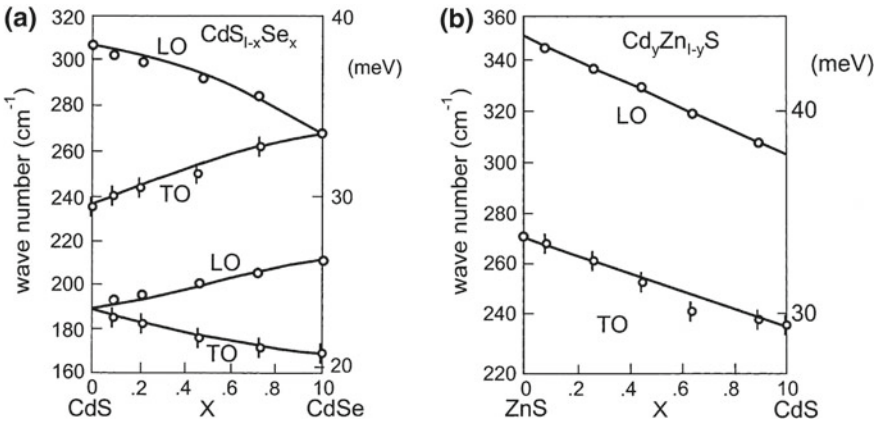


Fig. 13.1 The composition dependence of the optical phonon frequencies for a persistent mode type alloy, here $CdS_{1-x}Se_x$ (**a**) and an amalgamation type alloy, here $Zn_{1-y}Cd_yS$ (**b**). The data have been deduced from Raman and IR spectroscopy (adapted with permission from [68C1] ©1968 by American Physical Society https://doi.org/10.1103/PhysRev.172.924)

13.2 Defects and Localized Phonon Modes

In our discussion of phonons (and of our model substance in Chaps. 7 and 8), we assumed until now a perfect arrangement of the atoms on sites of the crystal structure. In fact, every real semiconductor contains many crystal or lattice defects. There are point defects, one-dimensional defects like dislocations, two-dimensional defects like stacking faults and (small angle) grain boundaries, or three-dimensional ones like precipitates or voids.

In the present context the most important are the *point defects*. For an introduction to this topic see, e.g., textbooks on solid state physics like [16A1] or [81L1, 83B1]. Point defects include vacancies, interstitial and substitutional atoms. We give some examples for point defects in a compound semiconductor *AB* in Fig. 13.2. We will come back to the electronic properties of semiconductors related to defects later. What we consider for the moment is the fact that a point defect may have a different mass m' and/or chemical binding, i.e., spring constant D' as compared to the atom which would be at this place in a perfect crystal.

A consequence of such a point defect is that a *localized phonon* mode may appear. This is a mode which cannot propagate through the sample with a plane-wave factor as in (10.13a) and (10.13b) or (10.18). Instead the amplitude has a maximum at the place of the defect and decays exponentially with increasing distance from it. Obviously such a mode is localized at or in the vicinity of the point defect.

If the eigenfrequency ω_{loc} of such a localized phonon mode falls into the bands of the intrinsic acoustic or optic modes and couples to them, it will not produce a big effect. Once such a localized mode is excited, it decays rapidly into bulk modes. The situation is different if ω_{loc} falls either in a spectral region where the pure material has no eigenfrequencies at all, or if the localized phonon couples only weakly to the bulk modes. Then the localized mode can produce, for example, an additional absorption band or Raman satellite. We return to this point later. The situation $\omega_{\mathrm{loc}} > \omega_{\mathrm{bulk}}^{\mathrm{LO}}$ can be realized for example when a substitutional atom is incorporated, which has

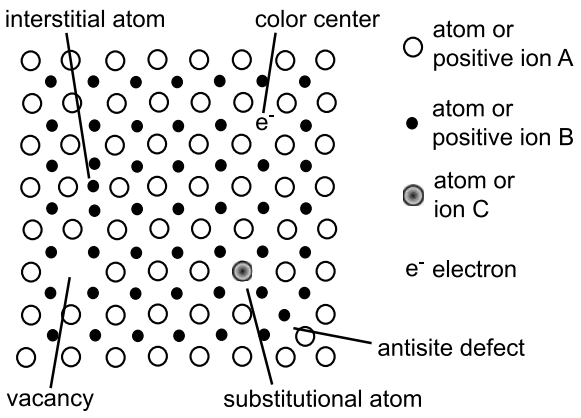


Fig. 13.2 The lattice of a semiconductor AB containing various types of point defects

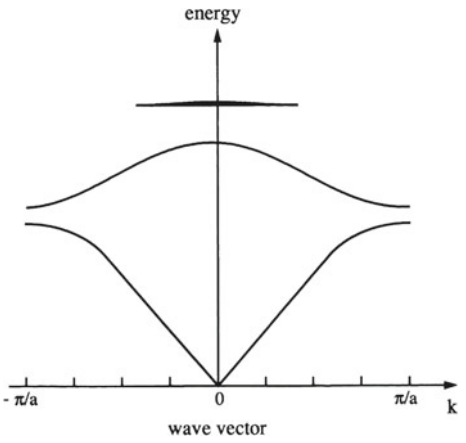


Fig. 13.3 A localized phonon mode represented schematically in a dispersion diagram

approximately the same coupling constant D but a much lighter mass m' than the atom which it replaces, according to (10.21) or Fig. 10.7.

Since the translation invariance of the lattice is destroyed at the point defect, the eigenfrequency ω_{loc} cannot be connected with a well-defined wave vector. There is, however, a possibility to incorporate a localized mode in a dispersion relation based on the following consideration. A localized mode can be constructed in a Fourier transformation-like method by a superposition of bulk modes with appropriate coefficients:

$$u_{\mathrm{loc}}(\boldsymbol{r}) = \sum_{\boldsymbol{k},i} a_{\boldsymbol{k},i} u_{\boldsymbol{k},i}(\boldsymbol{r}) \ , \tag{13.1}$$

where the index i runs over the various branches. Modes which are localized to one unit cell will need contributions from the whole Brillouin zone, while those which are more extended in real space involve contributions from smaller wave vectors only. We can now indicate a localized mode in the dispersion relation by a horizontal line covering the region of $\boldsymbol{k}$ values that make substantial contributions to the expansion of (13.1). We show such a situation in Fig. 13.3 where the thickness of the horizontal line is related to the modulus of the coefficients in (13.1).

As an example for the appearance of localized phonon modes we show in Fig. 13.4 an absorption spectrum of a GaAs crystal doped with Si. This dopant can occupy Ga or As sites as $\mathrm{Si_{Ga}}$ and $\mathrm{Si_{As}}$ and thereby act as donor or acceptor, respectively. Additionally the samples have been co-doped with Li to keep the concentration of free carriers low since the Si acceptors and donors do not compensate completely. The absorption spectrum of this system is shown for the energy range above the optical bulk phonon modes. Most of the peaks can be identified as local phonon modes of the centers or complexes indicated by arrows. Since Si and Li have considerably lower atomic masses than Ga or As and since the force constants are not too much lower, it is not surprising that the localized phonon modes appear at frequencies above the

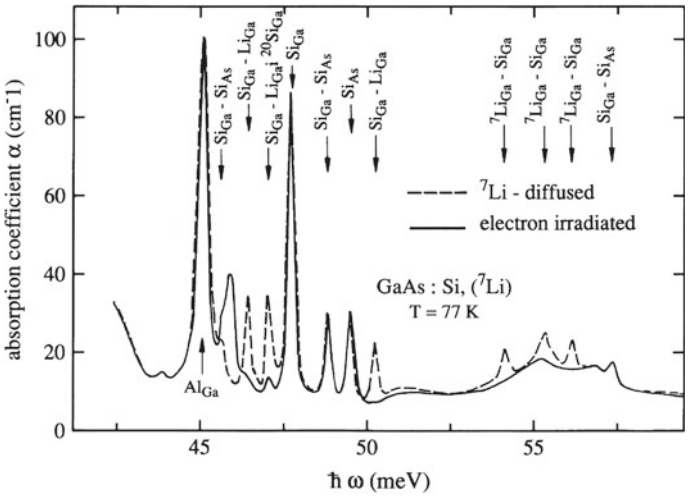


Fig. 13.4 Absorption spectrum of a GaAs:Si, Li doped sample, showing localized phonon modes (reprinted with permission from [71S1] ©1971 Springer Nature)

bulk modes. The isotope shift introduced by ^{7}Li and ^{8}Li is clearly visible in further data of [71S1]. More information on this topic may be found in [89J1, 94R1, 02K1].

Localized phonon modes can also be observed in Raman scattering or in luminescence and absorption in the visible as side-bands.

It should be noted that every defect is a scattering center for phonons and contributes, together with the anharmonicity mentioned above, to the finite phase relaxation time of phonons or to their finite contributions to the heat conductivity at low temperatures. If no scattering processes for phonons would exist at low temperature, their heat conductivity would diverge.

To conclude this discussion we would like to stress that even different isotopes act as scattering centers due to their different masses. If, for example, the intrinsic ballistic propagation of phonons is to be investigated, it is desirable to use materials which consist of one isotope only. For more information on this topic see, e.g., [81D1] or [94C1].

13.3 Phonons in Superlattices and Quantum-Well Structures

Until now we were mainly considering homogeneous, three-dimensional semiconductors. For the deduction of phonon properties, however, we started with a one-dimensional chain. Still we stated that the dimensionality does not have significant influence on the dispersion relation but it does on the density of states, according

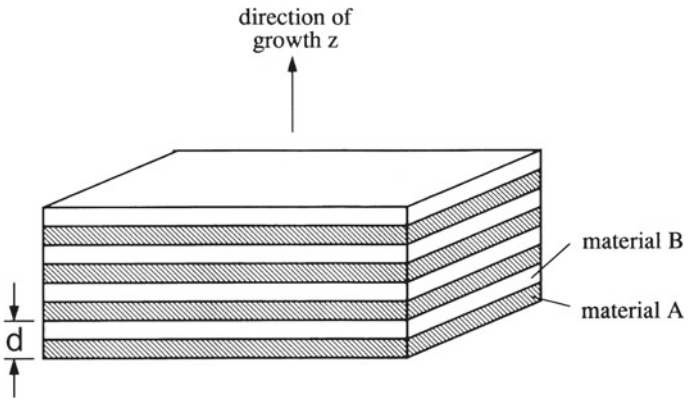


Fig. 13.5 A superlattice consisting of alternating thin layers of two different materials

to our discussion in Chap. 4. For the regime of a linear dispersion relation, i.e., for acoustic phonons with not too large wave vectors (see Figs. 10.11, 11.1, 11.2, 11.3) we get with (4.14)

$$D(\omega) \sim \omega^{d-1} \tag{13.2}$$

where d is the dimensionality of the system.

Now we want to address in the properties of phonons in so-called *superlattices (SLs)* which are periodic structures consisting of a large number of thin alternating layers of two different materials, as shown in Fig. 13.5. The different layers are only a few lattice constants thick and can be grown with atomic precision by various epitaxial techniques (see, e.g., [13P1]). Especially well-suited for growing superlattices are semiconductor materials with (nearly) matching lattice constants like the III–V compounds GaAs, AlAs and their alloys $Al_{1-y}Ga_yAs$. Also $Al_{1-x}In_xAs$ and $Ga_{1-y}In_yAs$ have matching lattice constants (also with InP as substrate) for certain compositions x,y. Superlattices from these materials have applications e.g., in quantum cascade lasers [03F1]. Superlattices have also been realized for other combinations of semiconductors like Si and $Si_{1-x}Ge_x$ or ZnSe and $ZnS_{1-x}Se_x$. Due to the differences in lattice constants the respective layers in these structures are strained (*strained layer superlattices*).

The formation of a superlattice results in strong modifications of the dispersion relation for phonons propagating in growth direction, i.e. normal to the planes, while the in-plane dispersion is generally at least not qualitatively modified. We concentrate therefore on the growth direction, which is frequently labeled the z-direction.

13.3.1 Back-Folded Acoustic Phonons

As can be seen from Fig. 13.5 a new artificial periodicity d occurs which is an integer multiple n_i of the (possibly strained) lattice constants a_i ($i = A, B$) of materials A and B

$$d = n_A a_A + n_B a_B \tag{13.3}$$

This is why such structures are called (artificial) superlattices.

Periodicity in the atom arrangement of the real lattice defines the range of valid $\boldsymbol{k}$-vectors in reciprocal space. Similar to the arguments used in connection with Fig. 10.10 we can now introduce a Brillouin zone. It extends in z-direction no longer from $-\pi/a \leq k_z \leq \pi/a$ but only from

$$-\frac{\pi}{d} \leq k_z \leq \frac{\pi}{d} \tag{13.4a}$$

All parts of the dispersion curve outside this interval can be shifted into this first so-called *mini-Brillouin zone* by vectors of the new reciprocal lattice

$$G_z = l_3 \frac{2\pi}{d}; \quad l_3 = 0, \pm 1, \pm 2, \ldots \tag{13.4b}$$

For a single acoustic phonon branch this procedure results in a dispersion relation like in Fig. 13.6, where we show the first few of these *back-folded acoustic phonon* branches.

The actual dispersion relation which enters is an average of the dispersion relations of the two materials, weighted by their relative thicknesses (the so-called Rytov-model) [89J1]. Usually one can observe only the few lowest, back-folded acoustic

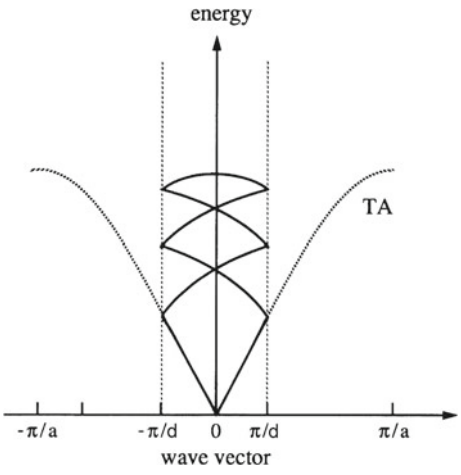


Fig. 13.6 Back-folded acoustic phonon branches in the mini-Brillouin zone of a superlattice of period d

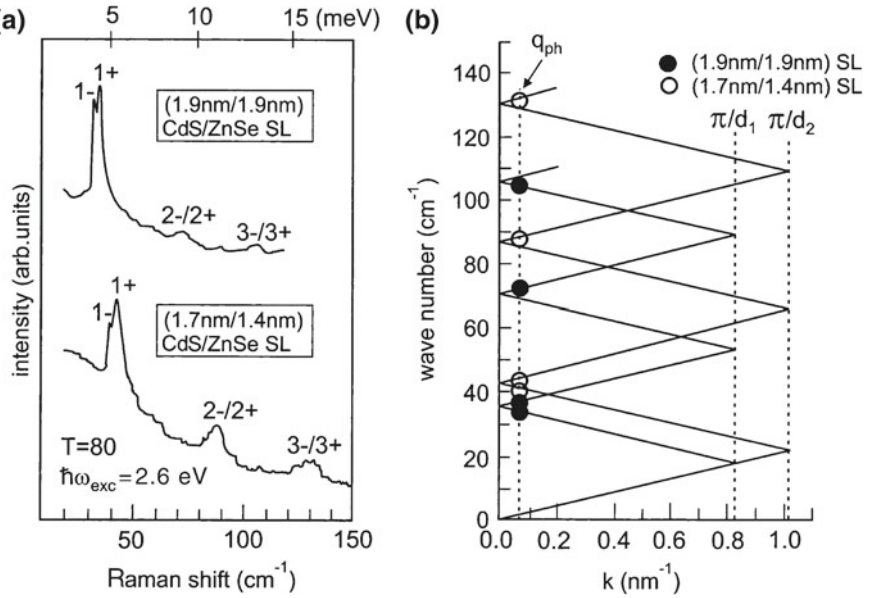


Fig. 13.7 Raman spectra of the first three back-folded longitudinal acoustic phonons (FLAPS) in two different CdS/ZnSe superlattices (**a**) and their calculated dispersion relation together with the data points from **a**, **b** (reprinted with permission from [01D1] ©2001 American Physical Society https://doi.org/10.1103/PhysRevB.64.245310)

phonon branches. The higher ones are broadened and washed out e.g., due to thickness fluctuations of the layers.

Raman spectra of CdS/ZnSe superlattices show the first three back-folded LA modes (so-called FLAPS) (see Fig. 13.7). The experimental data of different samples are in excellent agreement with calculated dispersion curves of the back-folded longitudinal acoustic phonons in the respective superlattice structures.

Under favorable conditions it is even possible to observe back-folded acoustic phonons directly in the IR. An example of a CdS/CdSe superlattice is given in [98G1]. Data for back-folded acoustic phonons in other systems like GaAs/AlAs or CdS/ZnS superlattices are found, e.g., in [72C1, 93R1, 99D1].

13.3.2 Confined Optical Phonons

For the optical phonon branches two different things can happen. If the branches in the two different materials overlap energetically, back-folded optical phonons result in a similar way as for the acoustic phonons. If they do not overlap energetically, TO or LO optical phonon oscillations in one material do not find a resonant partner in the other one (and vice versa) similarly to the case of alloys. Consequently,

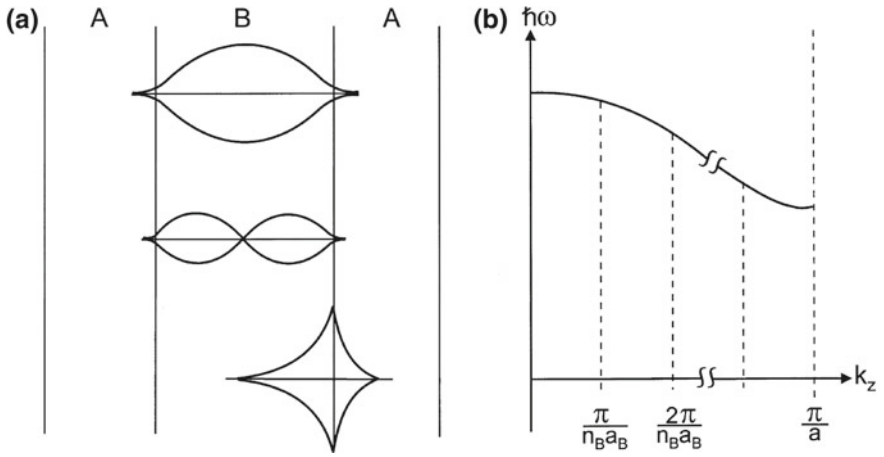


Fig. 13.8 Schematic representation of the envelope of an optical phonon mode confined to material B in a superlattice and of an interface phonon (**a**); the discrete values of k_z on an optical phonon branch of material B (**b**)

the phonons become localized or confined to the respective layer (*confined optical phonons*). The elongation-pattern of the atoms has nodes at the interface or extends only marginally with an exponentially decaying amplitude into the adjacent material as shown schematically in Fig. 13.8a. The same holds for confined phonons in single or multiple quantum well structures.

With minor corrections it can be stated that an integer number of half-waves has to fit into the layer of one material resulting in possible values for k_z in one layer of thickness $n_i a_i$

$$k_z^{(m)} = m\frac{\pi}{n_i a_i}; m = 1, 2, 3, \ldots; \quad i = A, B \tag{13.5}$$

These discrete k_z values are mapped in Fig. 13.8b on the dispersion relation of an optical phonon branch of e.g., material B. The observation of the higher orders of such confined optical phonons allows investigating the dispersion relation $\hbar\omega(k_z)$ over large parts of the first bulk Brillouin zone [89J1, 94C1, 94R1].

As an example we use again the CdS/ZnSe system which is a persistent-mode type. Consequently there are confined optical phonon modes in the two different sublattices. In Fig. 13.9a the displacement patterns of the first two confined LO phonon modes are shown for the CdS and for the ZnSe layers. Due to the mismatch of the eigenfrequencies, the amplitudes decay extremely fast beyond the boundary. The Raman spectrum shows well-resolved the first two confined LO phonon modes in ZnSe, while those in CdS are more broadened and the corresponding structure is washed out. The polarization-dependent selection rules are obvious, since the odd confined phonon modes are seen preferentially in off-diagonal polarization and the even ones in diagonal polarization. The calculated energies are marked by the dashed lines.

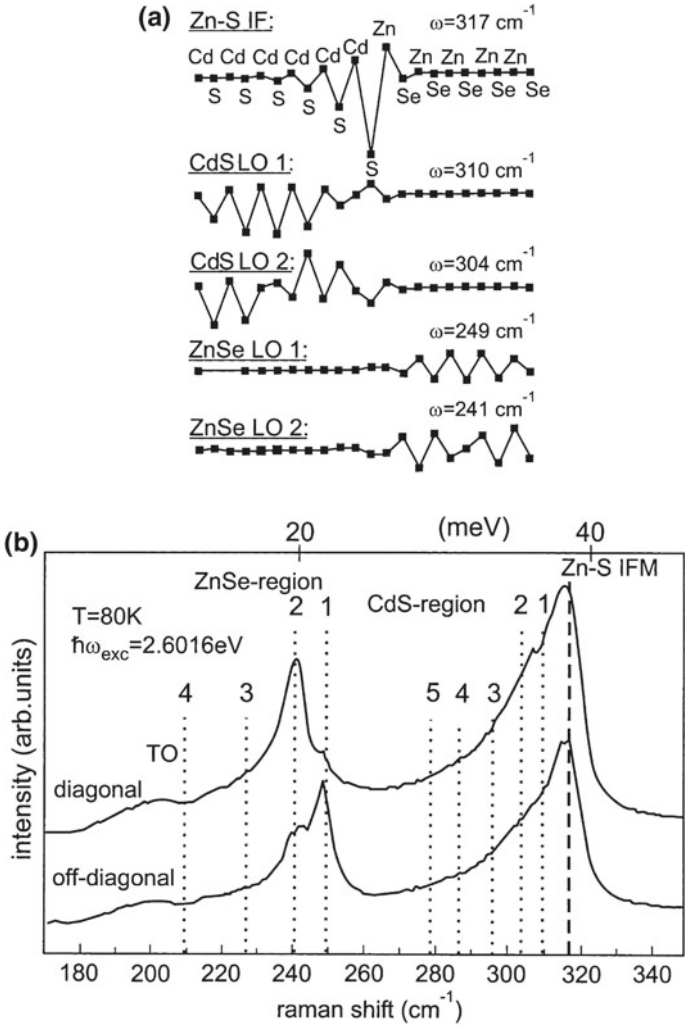


Fig. 13.9 Calculated atomic displacement patterns for the first two confined longitudinal optical phonon modes in a $(CdS)_6/(ZnSe)_5$ superlattice and of the ZnS interface mode (**a**); Raman spectra for parallel (*diagonal*) and orthogonal (*off-diagonal*) polarization with the calculated frequencies indicated by *dashed vertical lines* (**b**) (reprinted with permission from [01D1] ©2001 American Physical Society https://doi.org/10.1103/PhysRevB.64.245310)

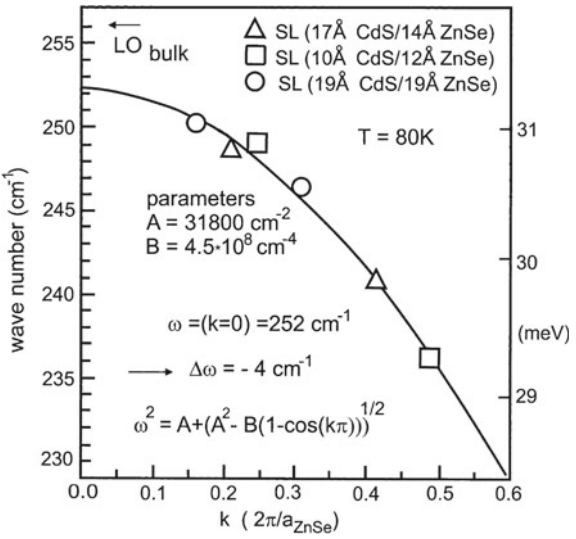


Fig. 13.10 The dispersion relation of LO phonons of ZnSe over one half of the first Brillouin zone deduced from confined optical phonons in a CdS/ZnSe superlattice (reprinted with permission from [01D1] ©2001 American Physical Society https://doi.org/10.1103/PhysRevB.64.245310)

By measuring the eigenfrequencies of confined phonons of different order and for superlattices of different layer thicknesses it is possible to investigate optically the phonon dispersion over a substantial part of the first Brillouin zone. This includes possibly some minor shifts or modifications in strained-layer superlattices due to the phononic deformation potentials. An example is shown for ZnSe in Fig. 13.10. The dispersion relation is shifted by $4\,\mathrm{cm}^{-1}$ due to strain as compared to bulk ZnSe. This technique is much easier and cheaper to realize than inelastic neutron scattering and can be performed in almost every well-equipped laboratory for optical spectroscopy. More examples of this type of experiments are found, e.g., in [93R1, 94C1, 97V1, 01L1, 10Y1].

13.3.3 Interface Phonons

If the superlattice is made of a ternary system, e.g., layers AB and AC, generally no new valences and oscillation modes will appear at the interface, which do not already exist in one of the two materials. The situation is different for quaternary systems of alternating layers AB and DC. In this case new modes may appear at the interface (here AC and BD oscillations) which exist in neither of the parent compounds. An example of a system of this type are CdS/ZnSe superlattices [01D1]: Depending on the chemical termination of the layers, there may be CdSe and ZnS modes. If these modes are situated energetically in a region, where no eigenmodes (confined or back-folded) of AB or DC (here of CdS or ZnSe) occur, these *interface phonon* modes can exist only at the interface with amplitudes decaying to both sides as shown

schematically in Fig. 13.8a. In this sense they correspond to the surface (polariton) modes of Sect. 8.6. A similar system is BeTe/ZnSe.

In the CdS/ZnSe system, actually two bonds can appear at the interface, which are not present in either of the two compounds, namely CdSe and ZnS bonds. CdSe interface phonon modes are difficult to detect since their frequency is rather low due to the heavy masses of Cd and Se. Also they might consequently merge in some other confined or back-folded modes. The situation is different for the ZnS mode, which involves the light cation and anion masses. The peak with the highest energy in the Raman spectrum of Fig. 13.9b is attributed to this interface mode. The displacement pattern is shown in Fig. 13.9a. The excellent coincidence of the calculated and measured energies supports this interpretation.

Interface phonons may also occur at the boundary between a polar material (e.g., GaAs) and a non-polar one (e.g., Si) (see e.g., [10Y1]).

We may summarize at this point that the artificial structure of superlattices introduces new phenomena like back-folded, confined and interface phonon modes. Such phonons have been studied in many superlattice and quantum-well structures. Please refer for details e.g., to [01L1]. For coherent phonons (back-folded acoustic or confined optical) in superlattices and quantum wells see [97M1, 08F1].

13.4 Phonons in Quantum Dots

The extreme case of phonon confinement in all three directions of space may occur in so-called *quantum dots* (also termed quantum boxes, nano-crystallites or nano-islands, depending partly on the way of their fabrication (see Appendix D). Here the dispersion relation of the phonons breaks up into a series of discrete modes, which is governed by the boundary conditions (e.g., free or clamped). For details see e.g., [92N1, 93T1, 95O1, 96W1, 98A1].

A peculiar consequence of the confinement is that acoustic phonons show a discrete energy spectrum starting at finite energy. There are modes where, e.g., one half of a spherical quantum dot rotates against the other and others where a sphere oscillates adopting the shape of an American football. This illustrates the importance of the boundary conditions given by the dot. The discrete acoustic modes can be observed, e.g., as satellite lines in resonant photoluminescence or photoluminescence excitation spectroscopy. The energies of the confined acoustic phonons shift according to R^{-1} where R is the dot radius and can thus be distinguished from other splitting mechanisms with a comparable order of magnitude but a different size dependence. E.g., the singlet-triplet splitting also observed in [96W1] varies with powers $R^{-2.5}$ to R^{-3}.

Due to the small size of quantum dot also an effective medium theory might be applicable. These theories have been tested e.g., for ZnS nano-particles or for SiO_2-Ge composite films in [78G1, 92F1, 06H1].

Problems

13.1. (a) Derive a simple formula to estimate the average sound velocity in a superlattice from two materials A and B with thicknesses d_A, d_B and speeds of sound v_A, v_B, respectively. (b) Try to estimate the average speed of sound and the respective speeds for CdS and ZnSe from the data of Fig. 13.7a.

13.2. Sketch the qualitative atomic displacement of the first few (small wave vector in growth direction) confined optical phonon modes in an $(GaAs)_n(AlAs)_n$ superlattice. What do you expect for the movement of the atoms right at the interface between the layers?

References

[68C1] I.F. Chang, S.S. Mitra, Phys. Rev. **172**, 924 (1968)
[71S1] W.G. Spitzer, Festkörperprobleme/Adv. Solid State Phys. **11**, 1 (1971)
[72C1] C. Colvard et al., Rev. Lett. **28**, 811 (1972)
[78G1] C.G. Granqvist, O. Hunderi, Phys. Rev. B **18**, 2897 (1978)
[81D1] B. Di Bartolo (ed.), *Nonlinear Spectroscopy of Solids: Advances and Application (1993)*. NATO ASI Series B, vol. 339 (Plenum Press, New York, 1994)
[81L1] M. Lannoo, *Point Defects in Semiconductors I: Theoretical Aspects*. Springer Series in Solid-State Sciences, vol. 22 (Springer, Berlin, 1981)
[83B1] J. Bourgoin, *Point Defects in Semiconductors II: Experimental Aspects*. Springer Series in Solid-State Sciences, vol. 33 (Springer, Berlin, 1983)
[89J1] B. Jusserand, M. Cardona, in *Topics in Applied Physics*, vol. 66 (Springer, Berlin, 1989)
[92F1] M. Fujii et al., Phys. Rev. B **46**, 15936 (1992)
[92N1] S. Nomara, T. Kobayashi, Solid State Commun. **82**, 335 (1992)
[93R1] T. Ruf et al., Phys. Rev. Lett. **71**, 3035 (1993)
[93T1] A. Tomaka, S. Onari, T. Arai, Phys. Rev. B **47**, 1237 (1993)
[94C1] M. Cardona, Festkörperprobleme/Adv. Solid State Phys. **34**, 35 (1994)
[94R1] T. Ruf et al., Festkörperprobleme/Adv. Solid State Phys. **34**, 237 (1994)
[95O1] S. Okamoto, Y. Masumot, J. Lumin. **64**, 253 (1995)
[96W1] U. Woggon et al., Phys. Rev. B **54**, 1506 (1996)
[97M1] R. Merlin, Solid State Commun. **102**, 207 (1997)
[97V1] R.V. Velasco, F. Garcia-Moliner, Surf. Sci. Rep. **28**, 123 (1997)
[98A1] A.P. Alivisatos et al., J. Chem. Phys. **89**, 4001 (1998)
[98G1] M. Göppert et al., Phys. Rev. B **57**, 13068 (1998)
[99D1] A. Dinger et al., Phys. Status Solidi (b) **215**, 413 (1999)
[01D1] A. Dinger et al., Phys. Rev. B **64**, 245310 (2001)
[02K1] A. Kaschner et al., Appl. Phys. Lett. **80**, 1909 (2002)
[01L1] C. Klingshirn (ed.), *Landolt–Börnstein*. New Series, Group III, vol. 34C, Parts 1 and 2 (Springer, Berlin, 2001 and 2004); Part 3, ed. by E. Kasper, C. Klingshirn (2007)
[03F1] J. Faist, C. Sirtori, D. Hofstetter, *Quantum Cascade Lasers* (Clarendon Press, Oxford, 2003)
[06A1] K.C. Agarwal et al., Phys. Rev. B **73**, 045211 (2006)
[06H1] J. Han et al., J. Phys. Chem. B **110**, 1989 (2006)
[08F1] M. Först, T. Dekorsy, in *Coherent Vibrational Spectroscopy*, ed. by S. de Silvestri, G. Cerullo, G. Lanzani (2008), p. 129

[09A1] S. Adachi, *Properties of Semiconductor Alloys: Group-IV, III-V and II-VI Semiconductors* (Wiley, Hoboken, 2009)
[10Y1] P.Y. Yu, M. Cardona, *Fundamentals of Semiconductors*, 4th edn. (Springer, Berlin, 2010)
[13P1] U.W. Pohl, *Epitaxy of Semiconductors: Introduction to Physical Principles* (Springer, Berlin, 2013)
[16A1] N.W. Ashcroft, N.D. Mermin, D. Wei, *Solid State Physics*, Revised edn. (Cengage Learning Asia, Singapore, 2016)

Part III
Electron States in Semiconductors

Part III of this book on semiconductor optics is intended to describe the properties of electrons in semiconductors. We start with the implications of periodic crystal potentials like Bloch waves, basic bandstructure, holes and the effective-mass concept. We then introduce the consequences of lattice deformations like strain and phonons on electronic states and electron dynamics. Details of the semiconductor bandstructure are theoretically deduced in particular by application of $\boldsymbol{k} \cdot \boldsymbol{p}$-theory. The specifics of the most common semiconductors like Si, Ge, GaAs, ZnO or GaN will be elaborated.

We will then derive the strength of optical transitions between electronic bands from perturbation theory and how optical spectroscopy can be used to determine critical points of the bandstructure like the bandgap. This is followed by the description of the statistical occupation of band states.

We then concentrate on structures of reduced dimensionality like quantum wells, superlattices, quantum wires and dots since they are not only of quantum-mechanical interest but they also have many device application. Of high topical interest are monolayer semiconductor structures due to their specific bandstructures.

Electronic states related to lattice defects and in particular resulting from doping as well as their signatures in optical experiment will be illustrated. We will then describe the localization of electrons in semiconductors with disordered lattice potentials. This part ends with the properties of excitations in carrier gases like plasmons and surface plasmons as well as optical properties related to correlation effects in carrier plasmas.

Chapter 14
Electrons in a Periodic Crystal Potential

Abstract The properties of electrons in crystalline solids are strongly determined by the periodic lattice potential. As a preparation for the discussion of the semiconductor bandstructure we derive first some general properties of electrons in crystalline lattices. We start with some necessary approximations by ignoring lattice vibrations (adiabatic approximation) and focus on a single electron in an averaged potential (one-electron approximation). We will account for the translational symmetry of lattices by introducing Bloch waves as solutions of the Schrödinger equation which includes a periodic potential. We illustrate the quasi-particle nature of electrons and holes and elaborate the origin and implications of the effective mass. Finally we describe some consequences of carrier-phonon coupling in (at least partially) ionic crystals.

14.1 Adiabatic Approximation

If we want to describe the physical properties of a semiconductor, all we have to do in principle, is to solve the Schrödinger equation for this problem. The *Hamiltonian of a semiconductor* consists of terms related to the ion cores, consisting of the nucleus and the tightly bound electrons in the inner shells as well as to the outer (or valence) electrons. The former have coordinates $\boldsymbol{R}_j$ and and masses M_j, the latter are described by and $\boldsymbol{r}_i$ and m_0. The Hamiltonian now reads (see e.g. [96M1, 09R1]):

$$H = -\frac{\hbar^2}{2}\sum_{j=1}^{M}\frac{1}{M_j}\triangle_{\boldsymbol{R}_j} - \frac{\hbar^2}{2m_0}\sum_{i=1}^{N}\triangle_{\boldsymbol{r}_i} + \frac{1}{4\pi\varepsilon_0}\times\left(\sum_{j>j'}\frac{e^2 Z_j Z_{j'}}{\left|\boldsymbol{R}_j-\boldsymbol{R}_{j'}\right|} + \sum_{i>i'}\frac{e^2}{|\boldsymbol{r}_i-\boldsymbol{r}_{i'}|} + \sum_{i,j}\frac{e^2 Z_j}{\left|\boldsymbol{R}_j-\boldsymbol{r}_i\right|}\right). \quad (14.1)$$

Here, Z_j is the effective charge of the ion core j and the indices j and i run over all M ion cores and N electrons, respectively.

The terms in the first row of (14.1) yield the kinetic energies of the ion cores and the electrons, respectively. The terms of the second row describe the interactions

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_14

among the ions, among the electrons, and between the ions and electrons. In case of an external field (like an applied magnetic field) one would have to add a respective term which is ignored for the moment.

We want to stress here that out of the four fundamental interactions so far known, namely strong, electromagnetic, weak, and gravitational interaction, only the electromagnetic one is of importance for all properties like binding, transport and the optical properties discussed in this book. Concerning the electromagnetic interaction we will restrict our discussion here mainly to the electric ones but will including also magnetic effects later when we describe, e.g., the Zeeman effect, diamagnetic shift, magnetic polarons, and magnons. The reason is that electric interactions are usually much stronger than magnetic ones, basically since electric interactions begin with monopole-monopole (i.e., Coulomb) interaction, whereas magnetic interactions start only with dipole-dipole interactions, due to the absence of magnetic monopoles (see (2.1a)). Magnetic interactions, however, do have a certain subtle importance, e.g., in (diluted) magnetic semiconductors or, beyond the scope of this book, in electron paramagnetic or nuclear magnetic resonance experiments.

The wavefunction solving (14.1) depends on all coordinates $\boldsymbol{R}_j$ and $\boldsymbol{r}_i$ including spins which are here not considered explicitly yet.

$$H\phi(\boldsymbol{r}_i, \boldsymbol{R}_j) = E\phi(\boldsymbol{r}_i, \boldsymbol{R}_j) \tag{14.2}$$

Since the indices j and i run from one to M and N, respectively, both count of the order of 10^{23} particles per cm^3 of semiconductor, it is obvious that there is at present no realistic chance of solving (14.1) and (14.2) though a proper solution would, in principle, contain all information about a given semiconductor. If we do not want to get stuck at this point we must use some approximations to simplify the Hamiltonian of (14.1).

The most important one is the so-called *adiabatic* or *Born–Oppenheimer approximation*. It starts from the fact that the mass of an ion core is three to five orders of magnitude heavier than the one of a free electron, i.e.,

$$M_j \simeq 1836 \cdot A_j m_0 \tag{14.3}$$

where A_j is the mass number of ion j. The electric forces that bind the outer electrons to the atom are comparable to the ones which bind neighboring atoms or ions and can be described for small elongations by parabolic potentials, i.e., by force constants β. Thus we can easily see, even from classical arguments, that the highest resonance frequencies Ω with which ions can oscillate are much lower than the corresponding values ω for electrons

$$\Omega \simeq \left(\beta M_j^{-1}\right)^{1/2} \ll \omega = \left(\beta m_0^{-1}\right)^{1/2} \ . \tag{14.4}$$

As a consequence, the electrons can practically instantaneously follow the motion of the ion cores, but not vice versa. This is the essence of the adiabatic approximation.

On this basis we can now separate $\phi(\boldsymbol{r}_i, \boldsymbol{R}_j)$ into a product of a wavefunction which depends only on the $\boldsymbol{R}_j$ and describes the motion of the ion cores, and another one which gives the wavefunction of the electron system depending on the momentary values of the $\boldsymbol{R}_j$. In a next step we will further assume that all ions are fixed at their equilibrium positions $\boldsymbol{R}_{j0}$ resulting finally in

$$\phi(\boldsymbol{r}_i, \boldsymbol{R}_j) = \phi(\boldsymbol{r}_i)\phi(\boldsymbol{R}_{j0}) \; . \tag{14.5}$$

The effects that result from the deviation of the ions from their equilibrium positions are treated in perturbation theory. The deviations can be described in terms of phonons, which we already introduced in Part II of this book where we also described optical properties related to these lattice vibrations. Also the electrons can interact with these phonons leading, e.g. to the formation of polarons (see Sect. 14.6), to a finite conductivity or to carrier relaxation within the bandstructure.

14.2 One-Electron Approximation and Bloch's Theorem

Following the adiabatic approximation we assume for the Hamiltonian in (14.1) that the atoms are fixed at their equilibrium positions, i.e., we ignore that phonons are excited. Then the sum over the potentials of all atoms forms a periodic potential for the electrons. But we are still left with a Hamiltonian for about 10^{23} interacting outer electrons per cubic centimeter, which should be properly treated in a many-particle formalism. Unfortunately it is extremely difficult to handle this approach. Instead one generally uses a so-called *one-electron approximation*. The idea is the following: One assumes that the periodically arranged atoms and all interaction potentials between electrons together form a periodic potential $V(\boldsymbol{r})$ with

$$V(\boldsymbol{r} + \boldsymbol{R}) = V(\boldsymbol{r}) \, , \tag{14.6}$$

where $\boldsymbol{R}$ is a translation vector of the lattice (see Sect. 9.1).

This is a valid approach especially for semiconductors or simple metals. In semiconductors there is only a low density (typically between 10^7 and 10^{19} cm^{-3}) of free electrons (or holes) that easily influence their mutual motion. In simple metals the density of free carriers is higher, but the density of those influencing their mutual interaction is limited to the fraction situated in an interval of width $\Delta E \approx 4k_\mathrm{B}T \ll E_\mathrm{F}$ energetically around the Fermi energy.

It must be noted that there are different theories to treat systems where the mutual interaction between carriers are essential. These are, e.g., the Hubbard model to describe the insulator-metal transition or the Fermi-liquid theory of metals. These topics are, however, beyond the scope of this book and we refer the interested reader to [09R1].

We return now to the approximation (14.6). One calculates the eigenstates for one electron of the corresponding Schrödinger equation and populates these states with

electrons according to Fermi–Dirac statistics until all electrons have been accommodated. The potential $V(\boldsymbol{r})$ should ideally be calculated in a self-consistent way e.g. by Hartree or Hartree–Fock approaches. But simpler and more feasible approaches are often used such as a periodic arrangement of screened Coulomb potentials or a "muffin-tin" potential.

We do not want to dwell too deep into bandstructure calculations and rather refer to the literature (e.g., [96M1, 09R1]). But the bandstructure details are crucial for the understanding of the optical properties of the electronic subsystem of semiconductors. Therefore we will at least give an overview on some basic strategies to derive the bandstructure of crystalline solids.

There are essentially two approaches to the problem of bandstructure calculations. In the first case one starts with free electrons, which have the simple parabolic dispersion relation in the non-relativistic regime

$$E(\boldsymbol{k}) = \frac{\boldsymbol{p}^2}{2m_0} = \frac{\hbar^2 \boldsymbol{k}^2}{2m_0} \tag{14.7}$$

shown in Fig. 14.1a by the dashed line. One introduces a weak periodic potential (see (14.6)) as a perturbation. These techniques include the *empty lattice*, the *nearly free electron (NFE)*, *augmented plane wave (APW)* and *orthogonalized plane wave (OPW)* approaches. In the last two cases, terms are added to make sure that the plane waves are orthogonal to the deeper atomic levels. These terms are treated as potentials and are known as *pseudo-potentials*. These pseudo potentials may be energy dependent.

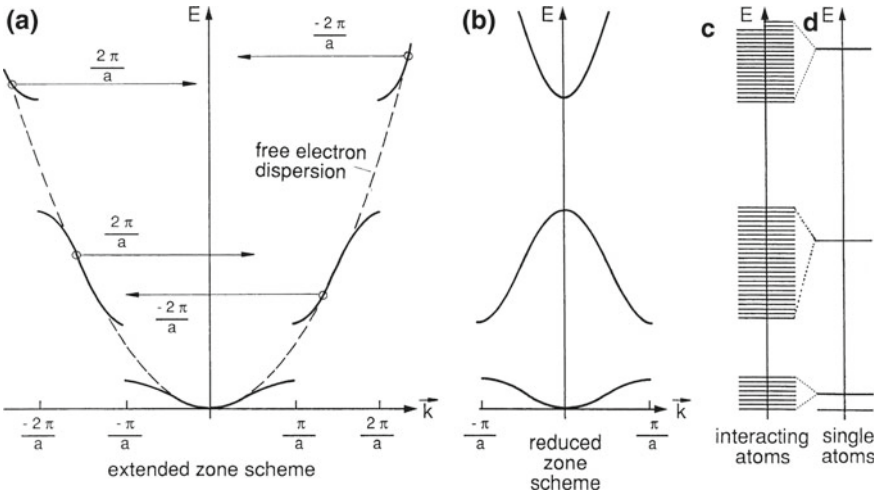


Fig. 14.1 The reduced-zone scheme (**b**) of electronic energy bands in a crystalline solid reached either starting from (nearly) free electrons (**a**) or from atomic orbitals (**c**, **d**)

The introduction of such a weak periodic potential does not alter the dispersion relation very much, but results in the formation of energy gaps at the borders of the Brillouin zones in which no stationary electron states exist (see Fig. 14.1a). We will explain this point in more detail in the following chapter.

The other group of methods starts with the atomic orbitals of the atoms forming the solid or more specifically the semiconductor. They involve the summation over one or more atomic orbitals placed at every atom site in the crystal and treat the interaction between the orbitals at neighboring sites, i.e., the wave function overlap as perturbation. These techniques of the *tight-binding approximation*, include the *linear combination of atomic orbitals (LCAO)* or the *extended Hückel method*.

The interaction between atoms leads to a splitting of the atomic level in as many states as there are atoms (Fig. 14.1c, d). This is similar to the scenario in classical mechanics where a weak coupling of two identical oscillators leads to two close-lying eigenfrequencies, three oscillators to three eigenfrequencies and 10^{23} oscillators to 10^{23} eigenfrequencies. The level spacing in the later case is so small that for all practical purposes the result is a continuous *energy band*. The first group of bandstructure calculation methods mentioned above is usually more suitable for *conduction-band* states of semiconductors the second one for *valence-band* states. The meaning of these two terms in the context of semiconductors will be illustrated in Sect. 14.3.

Independent of the individual approach, one finds in all cases that electron states in a periodic potential are energetically arranged in bands of a certain width, which may be separated by gaps in which no stationary eigenstates exist (Fig. 14.1). The electron eigenstates of a periodic potential are so-called *Bloch waves* $\phi_{\boldsymbol{k},i}(\boldsymbol{r})$ [28B1],

$$\phi_{\boldsymbol{k},i}(\boldsymbol{r}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\boldsymbol{r}} u_{\boldsymbol{k},i}(\boldsymbol{r}), \tag{14.8a}$$

where

$$u_{\boldsymbol{k},i}(\boldsymbol{r}) = u_{\boldsymbol{k},i}(\boldsymbol{r} + \boldsymbol{R}). \tag{14.8b}$$

The $\phi_{\boldsymbol{k},i}$ are evidently a product of a plane wave and a lattice-periodic term $u_{\boldsymbol{k},i}$ (14.8b), where $\boldsymbol{k}$ is the wave vector and i the index of the band, as shown in Fig. 14.2. The eigenenergies in the bands depend both on $\boldsymbol{k}$ and i and are periodic in $\boldsymbol{k}$-space, i.e.,

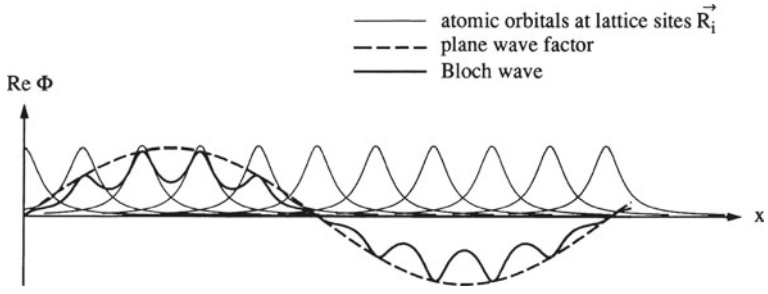


Fig. 14.2 Schematic drawing of the real (or imaginary) part of a Bloch wave in one dimension

$$E_i(\boldsymbol{k}) = E_i(\boldsymbol{k} + \boldsymbol{G}) \tag{14.9}$$

In a similar way one finds that

$$\phi_{\boldsymbol{k},i}(\boldsymbol{r}) = \phi_{\boldsymbol{k}+\boldsymbol{G},i}(\boldsymbol{r}). \tag{14.10}$$

These are the fundamental statements of *Bloch's theorem*. Since the $u_{\boldsymbol{k},i}(\boldsymbol{r})$ in (14.8a) derive from the wave function of the parent atoms and the exponential term reflects the plane wave character, the Bloch functions (14.8) incorporate the two limiting approaches (e.g., NFE and LCAO) mentioned above.

The Bloch theorem is also known as the *Ewald–Bloch theorem* since P.P.Ewald has found almost identical rules for the propagation of X-rays in crystals (see textbooks on solid state physics).

The statement of 14.8 once more allows the dispersion relation $E(\boldsymbol{k})$ to be reduced to the first Brillouin zone (Fig. 14.1a, b) in a similar way as was found for phonons in Sect. 10.2. Actually there are various possibilities to present the electronic bandstructure, which are shown in Fig. 14.3. Figure 14.3a shows the extended zone scheme in which various branches of the dispersion relation $E(\boldsymbol{k})$ are situated in various Brillouin zones. Equation (14.9) allows to shift the outer branches into the first Brillouin zone with suitable vectors of the reciprocal lattice $\boldsymbol{G}$ (here $\boldsymbol{G} = \pm 2\pi/a$) resulting in the reduced zone scheme of Fig. 14.1b or Fig. 8.3b. Alternatively we may repeat all branches periodically over the various Brillouin zones as shown in Fig. 14.3c.

Usually one uses the reduced scheme of Fig. 14.3b. The extrema of the various bands tend to occur either at the centre of the first Brillouin zone, i.e., at $\boldsymbol{k} = 0$ (the so-called Γ-*point*) or at the borders of the Brillouin zone as shown in Fig. 14.3b, but as we

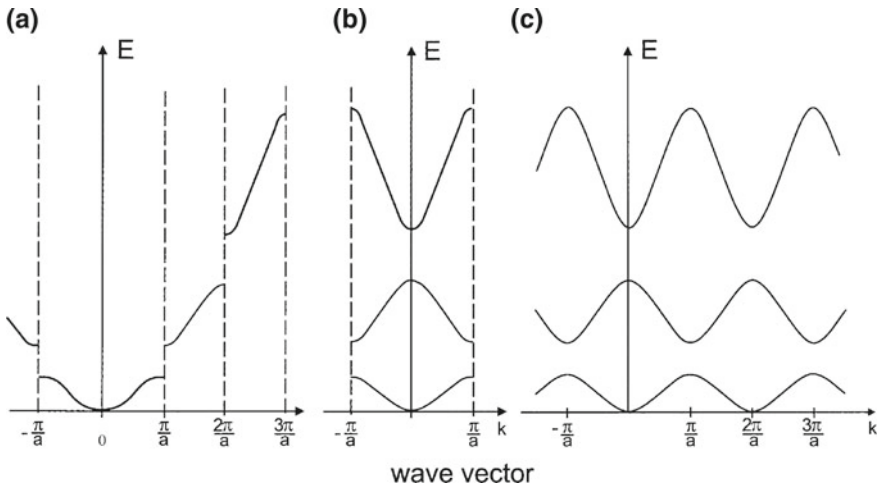


Fig. 14.3 The extended (**a**) the reduced (**b**) and the periodic representation of the electronic bandstructure (**c**) of a crystal

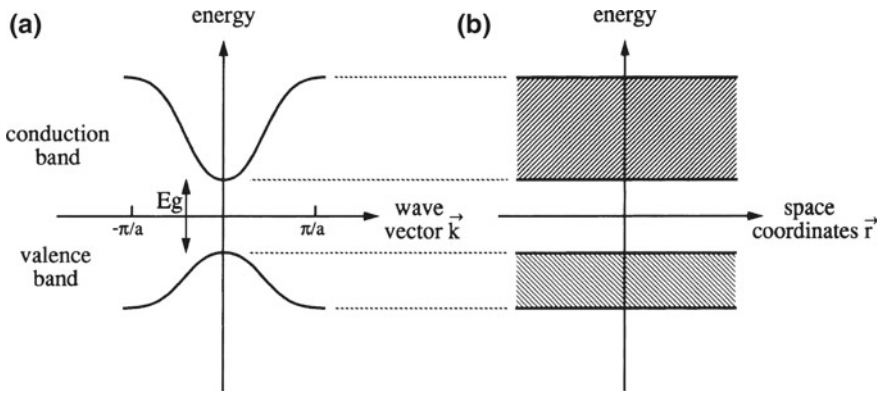


Fig. 14.4 The representation of bands in $\boldsymbol{k}$-space (**a**) and in real space (**b**)

shall see later, extrema may also occur somewhere else on the interval $0 \leq |k| < \pi/a$. The bands tend to be to a good approximation parabolic in the vicinity of the extrema but the positive or negative curvature can be different at every extremum and may even depend on the direction of $\boldsymbol{k}$ in systems of higher dimensionality than one including three. When approaching the border of the Brillouin zone, the slope of $E(\boldsymbol{k})$ normal to this border is generally zero.

With increasing energy (i.e. starting from the atomic $1s$ levels and going to the up-most occupied and the first unoccupied bands) the width of the allowed bands tends to increase, while the width of the forbidden gaps between them decreases or vanishes. It is even possible that various bands overlap energetically. We will give some arguments for this behavior later in this chapter.

To conclude this section we want to mention that the bands are usually represented in one of two different ways. The first is the dispersion relation (Figs. 14.1b, 14.3, and 14.4a). In the other case one plots the width of the bands and of the gaps between them as a function of the space coordinate $\boldsymbol{r}$ (Fig. 14.4b). The latter is especially useful to demonstrate spatial inhomogeneities or localized states. We shall meet some examples of both cases later.

14.3 Metals, Semiconductors, and Insulators

Having obtained a first insight into the electronic bandstructure in the above section, we now want to make the second step and put all electrons contributing to the solid by all the atoms into the bands using, of course, Fermi–Dirac statistics. We further assume zero temperature.

Some of the energetically lower-lying bands will be completely filled. We call bands which are completely filled at $T = 0\,\mathrm{K}$ *valence bands*, while all partly filled or empty bands are *conduction bands*. The important region for our purposes is that

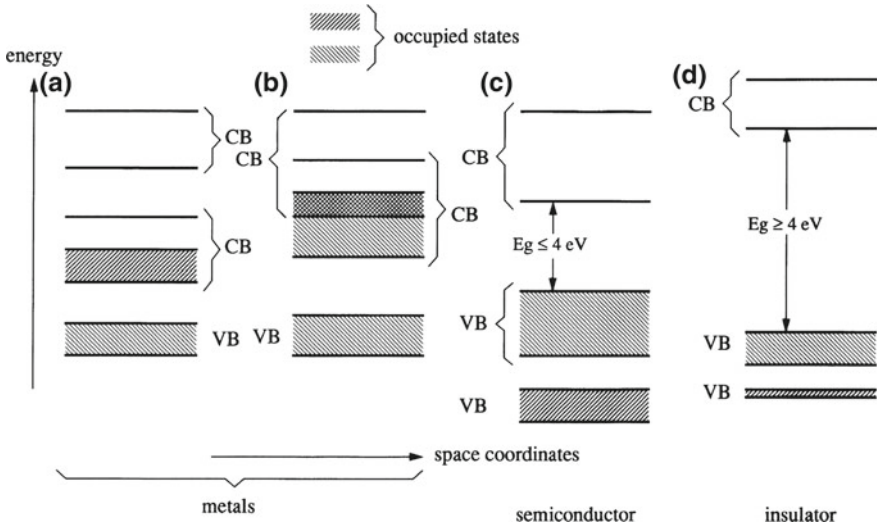


Fig. 14.5 Occupation of the bands for metals (**a**, **b**) a semiconductor (**c**) and an insulator (**d**)

around the highest valence and the lowest conduction bands. This region determines not only the optical properties around the fundamental absorption edge, but also the magnetic properties and the electronic contributions to the conductivity of electricity and of heat, respectively.

If the filling procedure of the bands ends in such a way that, at $T = 0\,\mathrm{K}$ there are one or more partly-filled conduction bands (Fig. 14.5a, b), we have a *metal* [28S1]. This situation arises for example if the atomic orbital which forms the band is itself only partly occupied (e.g., the outer s-level of the alkali metals Li, Na …) (Fig. 14.5a) or if a completely filled orbital forms a band which overlaps with a band stemming from an empty atomic orbital, as is the case in the rare earth metals (Ca, Mg …) in Fig. 14.5b.

If, on the other hand, the filling procedure gives one or more completely filled valence bands which are separated by a gap E_g from completely empty conduction bands, we have a *semiconductor* [31W1] for $0\,\mathrm{eV} < E_g \leq 4\,\mathrm{eV}$ and an *insulator* for $E_g \geq 4\,\mathrm{eV}$. The border line of 4 eV is set by convention and is not sharp. Diamond has, e.g., a gap of 5.5 eV, but is still considered to be a semiconductor, especially because it can be *n* or *p* doped.

Materials with $E_g = 0$ eV, i.e., materials in which the lowest conduction and uppermost valence band touch each other but do not overlap are called *semimetals*. Semiconductors with $0\,\mathrm{eV} < E_g \leq 0.5\,\mathrm{eV}$ are called *narrow-gap semiconductors*. In the range $0.5\,\mathrm{eV} < E_g \leq 2\,\mathrm{eV}$ we find the usual semiconductors, from which many are of wide technical importance like Ge, Si or GaAs. The regime $2\,\mathrm{eV} < E_g \leq 4\,\mathrm{eV}$ characterizes the *wide-gap semiconductors*, which lead continuously to the insulators. Please note, that there are some transition-metal compounds and rare-earth

oxides, which have partly occupied (deep lying) atomic d- or f-levels. Nevertheless these materials are insulators.

To conclude this section we want to introduce the terms *direct-gap* and *indirect-gap* semiconductor. A semiconductor is said to have a direct gap, or to be a *direct-gap semiconductor*, if the global maximum of the highest valence band and the global minimum of the lowest conduction band occur at the same point of the first Brillouin zone in the reduced zone scheme as in Fig. 14.4a. This point is frequently but not necessarily the Γ-point (i.e., $\boldsymbol{k} = 0$). There are also direct-gap semiconductors where the gap occurs at a specific point on the border of the Brillouin zone. Semiconductors, in which the two above-mentioned extrema occur at different $\boldsymbol{k}$ values in the Brillouin zone have an indirect gap or are called *indirect semiconductors*. The reason for this terminology comes from optical transitions between the band extrema. A photon with an energy equivalent to the typical width of a semiconductor band gap has (almost) vanishing momentum on the scale of the first Brillouin zone. Optical transitions are in this sense vertical in a $E(\boldsymbol{k})$ relation. If the band extrema occur at the same point in $\boldsymbol{k}$-space, a transition between the extrema is directly possible by absorption (or emission) of a photon. In the other case this transition is forbidden by the $\boldsymbol{k}$-conservation law and only indirect transitions are possible, which also involve the absorption or emission of a phonon for momentum conservation. We shall see examples for both cases in Sect. 15.2.

14.4 Quasi-particle Nature of Electrons and Holes in Crystalline Solids

The optical properties of the electronic system of semiconductors are largely determined by transitions of electrons between the upper valence bands and the lower conduction bands. The bandstructure of these electrons as presented so far, i.e., in connection with Figs. 14.1, 14.2, 14.3, 14.4 and 14.5, represents the so-called $N \pm 1$ *particle problem* in the following sense: if we consider a semiconductor with a completely filled valence band containing $N \simeq 10^{22} - 10^{23}\,\mathrm{cm}^{-3}$ electrons and a completely empty conduction band. Now we add one more electron and find that this electron has to be placed into a conduction band state. If we remove one of the N electrons and ask from which state it came, we find one of the valence-band states. An obvious task is now to describe the properties of one or of few electrons in an otherwise empty conduction band (CB). For an almost filled valence band, however, it is easier to consider the few empty states and their properties instead of the many occupied ones. This idea leads to the concept of *defect electrons* which are also called *holes*.

The properties of the hole are connected in the following way (Table 14.1) with the properties of the electron that has been removed from the valence band (VB): The hole has a positive charge and its wave vector and spin are opposite to those of the electron removed from the valence band. The two later statements are easy

Table 14.1 Properties of a hole in the valence band compared to the properties of the electron that has been removed from the valence band to create the hole

Property	Hole	Removed electron
Electric charge	q_h	$= -q_{re}$, $q_{re} \approx -1.6 \times 10^{-19}$As
Wave vector	$\boldsymbol{k}_h$	$= -\boldsymbol{k}_{re}$
Spin	σ_h	$= -\sigma_{re}$
Effective mass	$m_h > 0$	$= -m_{re}$, $m_{re} < 0$

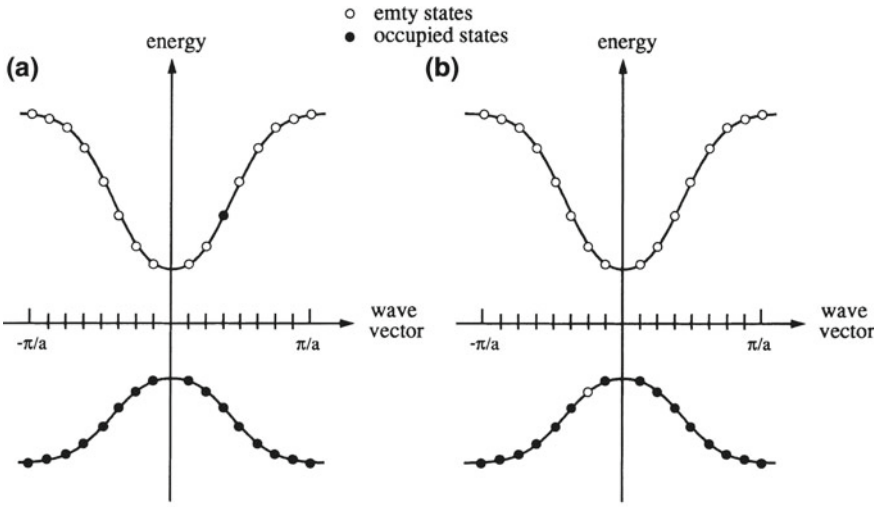


Fig. 14.6 One electron in the conduction band (**a**) and one hole in the valence band (**b**) representing the $N \pm 1$ particle problem. *Full circles*: occupied states; *open circles*: empty states

to understand. A semiconductor with a completely filled valence band has total momentum and spin both equal to zero. If we take one particle out, the remainder acquires for the above quantities values exactly opposite to those of the removed particle. For clarity Fig. 14.6a, b show the bandstructure containing one electron in the conduction band and one hole in the valence band, respectively. The states are equidistant in $\boldsymbol{k}$ (see Chap. 4) but we should note that there are usually 10^{22}–10^{23} states in each band per cm^3 and not only the few shown in Fig. 14.6.

The electrons and holes in a semiconductor crystal are *quasi-particles*. They can exist only in the crystal and not in vacuum, in contrast to normal electrons and positrons with which they have a lot in common, except the magnitude of the energy gap which is $\simeq$1 MeV for normal electrons and positrons i.e., twice the rest mass of $m_0c^2 = 511$ keV. The dispersion relations of electrons and holes are different from those of free electrons and positrons which for the non-relativistic case are given by

$$E_{\mathrm{e,p}} = \pm\left(m_0c^2 + \frac{\hbar^2k^2}{2m_0}\right), \tag{14.11}$$

where m_0 is the free electron mass.

The quantity $\hbar\boldsymbol{k}_{\mathrm{e,h}}$ of crystal electron and hole is a quasi-momentum, since it is conserved only modulo reciprocal lattice vectors (see (14.9) and (14.10)) and since the Bloch waves of (14.8) are not proper eigenstates of the momentum operator $\frac{\hbar}{\mathrm{i}}\nabla$.

Finally one should note that the energy of a hole increases if it is brought deeper into the valence band.

14.5 The Effective-Mass Concept

If we want to describe the motion of an electron or hole in a semiconductor under the influence of an external field (e.g., an electric or magnetic field) it is intuitively clear that we ought to consider a wave packet rather than the infinitely extended Bloch waves. To describe such a wave packet we superpose Bloch waves of a certain range of $\boldsymbol{k}$-vectors around a $\boldsymbol{k}_0$ as schematically given by:

$$\phi_{\boldsymbol{k}_0}(\boldsymbol{r}) = \sum_{\boldsymbol{k}} a_{\boldsymbol{k}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} u_{\boldsymbol{k}}(\boldsymbol{r})\,. \tag{14.12}$$

These types of wave packet are known as *Wannier functions*. In order to keep the $\boldsymbol{k}$-vector reasonably well defined, we localize the wavefunction only to a volume larger than a unit cell. Due to the uncertainty relation we would need wavefunctions from the whole Brillouin zone if we wanted to localize an electron to within one unit cell.

An external force $\boldsymbol{F}$ now changes the energy of the wave packet according to

$$\mathrm{d}E(\boldsymbol{k}) = \boldsymbol{F}\cdot\mathrm{d}\boldsymbol{s} = \boldsymbol{F}\cdot\boldsymbol{v}_{\mathrm{g}}\mathrm{d}t \tag{14.13}$$

with the group velocity given according to (2.18)

$$\boldsymbol{v}_{\mathrm{g}} = \frac{1}{\hbar}\mathrm{grad}_{\boldsymbol{k}}E(\boldsymbol{k}) \tag{14.14a}$$

or simplified

$$\boldsymbol{v}_{\mathrm{g}} = \frac{1}{\hbar}\frac{\mathrm{d}E}{\mathrm{d}k} \tag{14.14b}$$

Since, on the other hand, we have

$$\mathrm{d}E(\boldsymbol{k}) = \frac{\mathrm{d}E(\boldsymbol{k})}{\mathrm{d}k}\cdot\mathrm{d}\boldsymbol{k} = \hbar\boldsymbol{v}_{\mathrm{g}}\,\mathrm{d}\boldsymbol{k}, \tag{14.14c}$$

we find from combining (14.13) and (14.14c)

$$\hbar \frac{\mathrm{d}\boldsymbol{k}}{\mathrm{d}t} = \boldsymbol{F} = \dot{\boldsymbol{p}} \tag{14.14d}$$

This expression corresponds to Newton's law of motion, but now for the quasi-momenta $\hbar\boldsymbol{k}$ of the crystal electrons or holes.

Using (14.14d) we can describe in an abstract way the motion of a crystal electron when an electric field is applied to a semiconductor sample: a constant field evokes a constant force on the charge which leads to the fact that the quasi-momentum $\hbar\boldsymbol{k}$ changes by identical amounts for each (infinitesimally small) elapsed time interval. The single electron in the conduction band of Fig. 14.6 would progress along the states of the band until it reaches the edge of the Brillouin zone. It would then reenter the Brillouin zone from the opposite side and continue its progression within the band. This would happen periodically in time as long as the field is applied and would be accompanied by an oscillatory motion of the electron in real space (note that the quasi-momentum $\hbar\boldsymbol{k}$ changes its sign). This periodic motion of the crystal electron is called *Bloch oscillation*.

Actually this oscillatory behavior does not occur in real semiconductors since the carriers are scattered on a (sub-)picosecond time scale by phonons or defects. This leads only to a small change in average quasi-momentum of an ensemble of carriers and its drift with a constant drift velocity (averaged over all group velocities of the carries in the ensemble) in the electric field. Bloch oscillations can, however, be observed in superlattices, where the size of the Brillouin zone is drastically reduced by zone folding. We will come back to this oscillation dynamics in second volume of this book.

The discussion of the Bloch oscillations shows that electrons (and holes) in a crystalline solid react quite differently on applied forces as free carrier do. They are not simply accelerated in a certain direction—they might be decelerated or even change direction of motion after some time. Having this in mind, we still want to analyze the acceleration $\boldsymbol{a}$ of the wave packet (14.12) by a force by combining (14.14b) and (14.14d):

$$\boldsymbol{a} = \frac{\mathrm{d}\boldsymbol{v}_\mathrm{g}}{\mathrm{d}t} = \frac{1}{\hbar}\frac{\partial^2 E}{\partial k \partial t} = \frac{1}{\hbar}\frac{\partial^2 E}{\partial k^2}\frac{\mathrm{d}\boldsymbol{k}}{\mathrm{d}t} = \frac{1}{\hbar^2}\frac{\partial^2 E}{\partial k^2}\boldsymbol{F}\,. \tag{14.14e}$$

Comparing with the trivial form of (14.14d)

$$\boldsymbol{a} = \frac{1}{m}\boldsymbol{F} \tag{14.14f}$$

we find that the crystal electron and hole move under the influence of external fields through the crystal like particles, however, with an *effective mass* given by

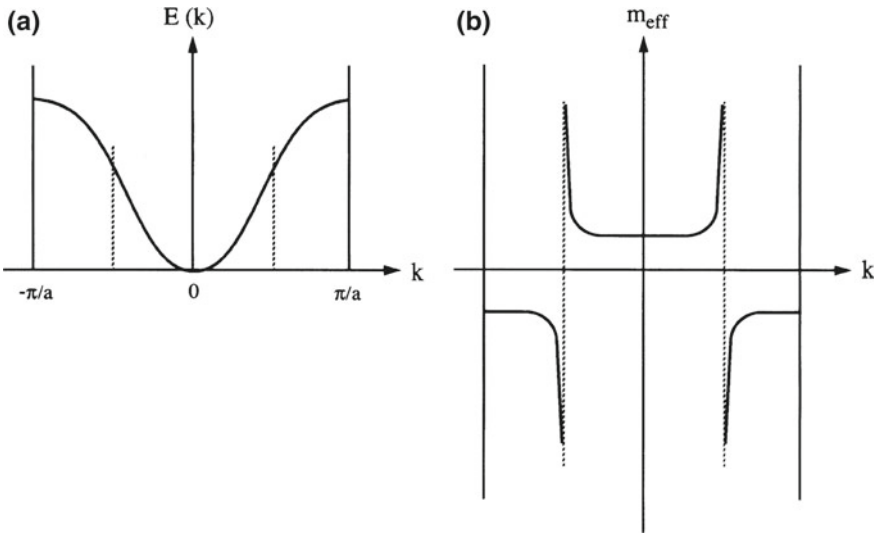


Fig. 14.7 Schematic dispersion (**a**) of a conduction band and the resulting effective mass (**b**) (adapted with permission from [09I1] ©2009 Springer Nature)

$$\frac{1}{m_{\text{eff}}} = \frac{1}{\hbar^2}\frac{\partial^2 E}{\partial k^2} = \frac{1}{\hbar^2}\frac{\partial^2 E}{\partial k_i \partial k_j}\,; \quad i, j = x, y, z\,. \tag{14.15}$$

From (14.15) we see that the inverse of the effective mass is given by the curvature of the dispersion relation. For free particles like electrons and positrons, which have a parabolic dispersion, this curvature and thus the effective mass is simply given by the rest mass $m_{\text{eff}} = m_0$. For electrons or holes in a crystal the situation is different: The curvature of a band can be positive or negative (see Fig. 14.6). Further the right-hand side of (14.15) shows that the effective mass is actually a tensor and can depend on the direction in which the electron or hole moves.

The important point now is that the bands of semiconductors tend to be parabolic in the vicinity of the band extrema, as shown schematically in Fig. 14.7 or in the real bandstructures discussed below (Figs. 15.7, 15.8, 15.9, 15.10, 15.11 and 15.12). These extrema are the most important parts of the bandstructure and determine the optical and transport properties. Since the effective masses are constant in these regions one can often use the so-called *effective-mass approximation*: Electrons and holes in a semiconductor are simply treated as free particles, but with an effective mass given by (14.15).

Please note that the mass of an electron is positive if the curvature of the band is positive. Due to the change of the sign of the properties of holes compared to those of the missing electron in the valence band (Table 14.1) the mass of the hole is positive at the maximum of the valence band where the curvature is actually negative.

The rationale behind the effective masses of electrons and holes is basically the following. The forces which act on a crystal electron are the ones from all the ions

and the other electrons in the crystal as well as the externally applied ones. For simplicity, we condensed the first two types of forces to yield the mean periodic potential of (14.1) using the one-electron approximation. Then we need to consider only the external forces acting on the electrons and the price we have to pay is the fact that the electrons and holes react with an effective mass. This mass is fortunately, in some regions at least, constant, allowing the effective-mass approximation mentioned above. But it can change as a function of $\boldsymbol{k}$ and become negative or even infinite, as can be seen in Fig. 14.7.

It is important to note that an increasing curvature of a band is necessarily connected with an increasing width of the band. Therefore we find the qualitative relation: electron easy to move and accelerate $\longleftrightarrow$ low effective mass $\longleftrightarrow$ large curvature of the band $\longleftrightarrow$ large band width $\longleftrightarrow$ strong coupling between adjacent atoms and vice versa. Thus we have here another example of the more general discussion given in the introduction to Sect. 8.4.

Furthermore it can be stated as a trend that a large band width leaves less space on the energy scale for gaps. Therefore we can also state as a rule of thumb, that narrow-gap semiconductors tend to have small effective masses and wide-gap semiconductors tend to have carriers with larger effective masses.

In the rest of this book we will use the effective-mass approximation if not stated otherwise. We will further omit the prefix "crystal" when talking about electrons and holes in semiconductors and we indicate the effective masses by indices e and h: m_e = effective mass of electron and m_h = effective mass of hole, sometimes with an additional index to distinguish different bands. But we always remember, that these charge carriers are quasi-particles, characterized by their band index, energy and (quasi-!)momentum.

It should be mentioned that the Fermi energy is often situated in metals with partly-filled bands in regions where the bandstructure $E(\boldsymbol{k})$ is no longer parabolic. On the other hand, the electrons close to the Fermi energy are in metals the relevant ones for many physical properties like the electrical conductivity, the specific heat or the paramagnetic behavior of the electron gas. Therefore the use of the effective-mass concept given by (14.15) is less frequently used in metal physics. It is replaced by another concept that is based on the fact that the first derivative $\mathrm{grad}_{\boldsymbol{k}} E(\boldsymbol{k})$ vanishes per definition close to the extrema of the conduction and valence bands of semiconductors but usually not around E_F in a metal. One can then define a *momentum effective mass* m_i [00S1] in contrast to the *force effective mass* m_eff in (14.14) and (14.15) via

$$\hbar \boldsymbol{k} = \boldsymbol{p} = m\boldsymbol{v} \tag{14.16}$$

resulting with $\boldsymbol{v}_\mathrm{g} = \frac{1}{\hbar}\mathrm{grad}_{\boldsymbol{k}} E(\boldsymbol{k})$ in the vicinity of $E_\mathrm{F}(\boldsymbol{k}_\mathrm{F})$ in

$$\frac{1}{m_i} = \frac{1}{\hbar^2 k_\mathrm{F}} \mathrm{grad}_{\boldsymbol{k}} E(\boldsymbol{k}=\boldsymbol{k}_\mathrm{F})\,. \tag{14.17a}$$

For the isotropic case this can be simplified to

$$m_i = \frac{\hbar^2 k}{\partial E(k)/\partial k} \tag{14.17b}$$

which is nothing else than the relation

$$m_i = \text{momentum}/\text{group velocity}\,. \tag{14.17c}$$

The effective masses appear also in the *mobility* of electrons and holes $\mu_{e,h}$ and in the *electrical conductivity* σ via

$$\sigma = ne\,\mu_e + pe\,\mu_h = ne^2\tau_e/m_e + pe^2\tau_h/m_h \tag{14.18}$$

with

$$\mu_{e,h} = e\tau_{e,h}/m_{e,h} \tag{14.19}$$

where n and p are the concentrations of electrons and holes and $\tau_{e,h}$ are the relaxation times. The latter are essentially given by the times between collisions of the carriers with, e.g., phonons or (charged) defects, which destroy the memory of the momentum of the carrier before the scattering. These times are also known as times of free flight of the carriers or collision times. The last term is rather misleading, because it does not mean the time a collision takes, but the time between two collisions.

We can conclude at this point that the effective-mass concept is extremely useful to describe the dynamics of electrons and holes in semiconductors.

14.6 Electron–Phonon Interaction and Consequences of Lattice Deformation

In the adiabatic approximation of Sect. 14.1 we eliminated the electron–phonon interaction to establish some general properties of crystal electrons. Before proceeding to some basic concepts of bandstructure calculations and to the bandstructures of real semiconductors, we want to discuss various consequences of electron–phonon interaction (for more details see [09R1]).

If we consider a free electron or a hole in a semiconductor which has at least partially ionic binding, the charge carrier will locally polarize the lattice provided we relax the assumption that the ions are fixed at their lattice sites. An electron will attract ions with a positive charge and repel those with a negative one (Fig. 14.8). For holes, the situation is just the opposite. We can describe this lattice distortion as a superposition of preferentially longitudinal optical phonons, i.e., a free carrier is accompanied by a "phonon cloud". The entity of charge carrier plus phonon cloud is called a *polaron*. In semiconductors, the radius of the phonon cloud is larger than the lattice constant, resulting in a so-called *large polaron* in contrast to the *small polaron*

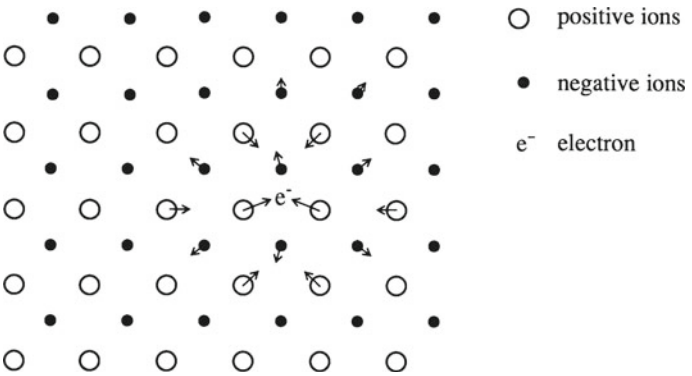


Fig. 14.8 The lattice distortion around a carrier in a (partly) ionic semiconductor illustrating the polaron concept

which occurs in ionic insulators and may lead to self-localization of a carrier within a unit cell [72B1].

The effective mass $m_{e,h}$ of a polaron is larger than that of an electron in a rigid lattice $m_{e,h}^{r.l.}$ [09R1]:

$$m_{e,h} \simeq m_{e,h}^{r.l.} \left(1 + \frac{\alpha_{e,h}}{6}\right) , \tag{14.20a}$$

where the index r.l. stands for "rigid lattice" and α is a dimensionless quantity which describes the *Fröhlich coupling* of carriers to the LO phonons. One finds

$$\alpha_{e,h} = \frac{e^2}{8\pi\varepsilon_0 \hbar\omega_{LO}} \left(\frac{2m_{e,h}^{r.l.}\omega_{LO}}{\hbar}\right)^{1/2} \left(\frac{1}{\varepsilon_b} - \frac{1}{\varepsilon_s}\right) . \tag{14.20b}$$

where ε_s and ε_b are the low- and high-frequency values of the dielectric function below and above the optical-phonon resonances, respectively (see also Chaps. 7 and 12). Obviously semiconductors without an ionic binding contribution have $\varepsilon_b = \varepsilon_s$ (see (7.26), and (7.28)) and thus $\alpha = 0$. For all normal semiconductors one finds $0 \leq \alpha \leq 1$.

Additionally the lattice relaxation leads to a decrease of the width of the bandgap by an amount ΔE_g with contributions for electrons and holes

$$\Delta E_g^{e,h} = \alpha_{e,h} \hbar\omega_{LO} . \tag{14.20c}$$

The radius of the phonon cloud in the polaron r_p is finally given by

$$r_p^{e,h} = \left(\hbar/2m_{e,h}^{r.l.}\omega_{LO}\right)^{1/2} . \tag{14.20d}$$

Basically it is very difficult to "fix the atoms at their lattice sites". Therefore all common experimental techniques to determine the effective masses of electrons and holes, e.g., by cyclotron resonance, or the value of the bandgap, e.g., by optical spectroscopy, will give polaron values. We therefore continue to use the effective-mass approximation and simply bear in mind that all values given for $m_{e,h}$ or E_g are actually polaron values.

In the so-called semi-magnetic semiconductors, usually II–VI compounds in which the cations are partly replaced by Mn or Fe ions, one finds so-called *magnetic polarons*. In this case the carrier spin results in a spin alignment of the paramagnetic ions in its vicinity. We will come back to this point in Sect. 24.1.

Another consequence of the interaction of the carriers with lattice deformation is the fact, that there is a general trend of a band-gap shrinkage in semiconductors with increasing temperature. The decrease $E_g(T=0) - E_g(T) = \Delta E_g(T)$ tends to vary quadratically with temperature at lower temperatures ($T \leq 100$ K) and linearly above. This behavior is often described by the Varshni formula [67V1]

$$\Delta E_g(T) = \frac{\alpha T^2}{\beta + \gamma T}. \tag{14.21}$$

which, however, shows some deviations at low temperatures. A better fit formula was given in [99P1]

$$\Delta E_g(T) = \frac{\alpha \theta_p}{2} \left[\sqrt[P]{1 + \left(\frac{2T}{\Theta_p}\right)^P} - 1 \right] \tag{14.22}$$

with Θ_p being an effective phonon temperature, α a high temperature limit of the entropy and the fractional exponent P a material specific parameter.

There are two main contributions to $\Delta E_g(T)$ namely the carrier-phonon coupling, which depends on the phonon occupation, and the temperature dependence of the lattice constant, since the bandgap depends on this quantity via the deformation potentials introduced below with (14.24). High-precision experiments on GaAs showed, that changes in the electron-phonon interaction dominate at low temperatures while both contributions contribute about equally at high temperatures [00L1].

Some semiconductors like CuCl or some lead salts actually show an increase in E_g with increasing temperature, for some others like CuBr $E_g(T)$ goes through a maximum with increasing temperature. For data see, e.g., [82L1] and references therein.

Semiconductors are sometimes grown epitaxially on substrates with a mismatch in the respective lattice constants. This leads to tensile or compressive strain in the semiconductor layer (see Appendix D). Strain can also be induced by application of pressure to a sample. Such strain has of course consequences for the electronic bandstructure. The width of the bandgap depends on strain via the deformation potential (see (14.24)), and the heavy and light-hole bands (see Sect. 15.2) split under the influence of strain already at $\boldsymbol{k} = 0$. In strained low-dimensional structures like

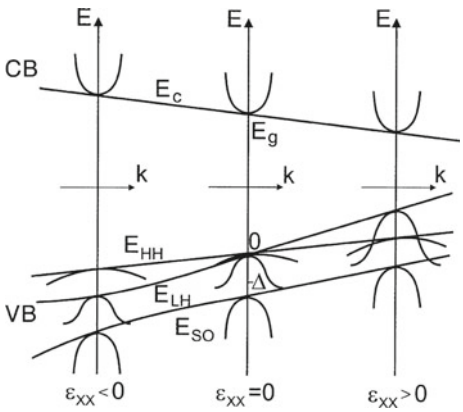


Fig. 14.9 Schematic drawing of the *lhh–lh* splitting, of the spin-orbit splitting and of the change of the bandgap under two-dimensional compressive ($\varepsilon_{xx} < 0$) or tensile strain ($\varepsilon_{xx} > 0$) (reprinted with permission from [93T1] ©1993 Elsevier https://www.sciencedirect.com/journal/physica-b-condensed-matter)

strained quantum wells or superlattices this effect comes in addition to the splitting of these bands by quantization as discussed in Sect. 17.2. The usual behavior under tensile and compressive strain is shown in Fig. 14.9.

Apart from the renormalization of mass and energy there are other consequences of carrier-phonon interaction, the most important being scattering phenomena. An electron or hole can be scattered inelastically, e.g., by emitting or absorbing a phonon under energy and momentum conservation between initial (i) and final (f) states:

$$E_{\mathrm{e}}^{\mathrm{i}} = E_{\mathrm{e}}^{\mathrm{f}} \pm \hbar\omega_{\mathrm{Phonon}}\,, \tag{14.23a}$$

$$\boldsymbol{k}_{\mathrm{e}}^{\mathrm{i}} = \boldsymbol{k}_{\mathrm{e}}^{\mathrm{f}} \pm \boldsymbol{k}_{\mathrm{Phonon}}(+\boldsymbol{G}). \tag{14.23b}$$

The interaction Hamiltonian which describes such processes can have various origins:

- Optical phonons are often accompanied by an electric field (Sects. 10.2, 10.3 and Chap. 12). The interaction of carriers with the electric field of preferentially longitudinal optical phonons is known as the *Fröhlich interaction*.
- Since the width of the gap depends on the lattice constant and on the arrangement of the atoms in the basis, a change of these quantities will influence the bandstructure. On the other hand, a phonon can be considered as a periodic deformation of the arrangement of atoms, and the carriers "feel" the resulting modulation of the bands. The resulting interaction between carriers and phonons is called *deformation-potential scattering*. The deformation-potential scattering occurs for both acoustic and optical phonons. The deformation potential Ξ connects usually the relative displacement $\Delta a/a$ of an atom with lattice constant a with the energetic shift of a band ΔE resulting schematically in

$$\Delta E = \Xi \frac{\Delta a}{a} \tag{14.24}$$

Typical values of deformation potentials are in the range of 5–10 eV. Note that (14.24) is valid only for $\Delta a/a \ll 1$.

- Finally, it is known that many non-centrosymmetric crystals show the piezoelectric effect, i.e., the appearance of an electric field as a consequence of strain, i.e., of lattice distortion. Again we can consider an acoustic or optical phonon as a periodic modulation of the lattice parameters, which produces, via the piezoelectric effect, a varying electric field which interacts with the electrons and holes. This effect is the so-called *piezo (acoustic) coupling*.

The discussion in this section demonstrated that the interactions of carriers with phonons and lattice deformations have some pronounced consequences. The experimentally deduced effective masses and band-gap energies, which will be used throughout this book, explicitly include these corrections. The scattering processes largely determine the dynamics of carriers, excitons and other excitations in semiconductors. We will frequently elaborate on these effects below.

Problems

14.1 The explicit treatment of the semiconductor Hamiltonian requires several approximations which finally lead to the one-electron approximation. Get some information on the various steps by reference to a text book on solid-state theory like [72Z1, 96M1, 09R1].

14.2 The effective mass of the electron in the simple conduction band of Fig. 14.7 diverges at the inflection points of the dispersion relation. How does an electron in such a state react on an electric field? Retrace the motion of a single electron within the band which is described in connection with the Bloch oscillations.

14.3 Show by application of the momentum operator to the respective wavefunctions of electrons in vacuum and Bloch waves, that the latter are no proper eigenstates of $\frac{\hbar}{\mathrm{i}}\nabla$.

14.4 Use a compilation of semiconductor data to find the highest and lowest values for effective electron and hole masses in group-IV materials and in binary III–V, II–VI and I–VII semiconductors. What is the trend of the dependence of the effective masses on E_g?

14.5 Why are the transition metals and rare earth oxides mentioned in Sect. 14.3, which have partly occupied d- or f-states, insulators?

References

[28B1] F. Bloch, Z. Phys. **52**, 555 (1928); ibid **59**, 208 (1930)
[28S1] A. Sommerfeld, Z. Phys. **47**, 1 (1928)
[31W1] A.H. Wilson, Proc. R. Soc. Lond. Ser. A **133**(822), 458 (1931). This is an early example for the use of the term "semi-conductor" and gives a good list of references from the time, when the basics of electronic properties of solids in general and of semiconductors have been developed
[67V1] Y. Varshni, Physica **34**, 149 (1967)
[72B1] R. von Baltz, U. Birkholz, Festkörperprobleme/Adv. Solid State Phys. **XII**, 233 (1972)
[72Z1] J.M. Ziman, *Principles of the Theory of Solids* (Cambridge University Press, Cambridge, 1972)
[82L1] O. Madelung, U. Rössler (eds.), *Landolt–Börnstein*. New Series, Group III, vol. 17a to i, 22a and b, 41A to D (Springer, Berlin, 1982–2001)
[93T1] T. Taguchi, Y. Kawakami, Y. Yamada, Phys. B **191**, 23 (1993)
[96M1] O. Madelung, *Introduction to Solid-State Theory*. Springer Series in Solid-State Sciences, vol. 2 (Springer Science & Business Media, New York, 1996)
[99P1] R. Pässler, Phys. Status Solidi B **216**, 975 (1999)
[00L1] D. Lüerßen, R. Bleher, H. Kalt, Phys. Rev. B **61**, 15812 (2000)
[00S1] C. Skierbiszewski et al., Appl. Phys. Lett. **76**, 2409 (2000)
[09I1] H. Ibach, H. Lüth, *Solid State Physics* (Springer, Berlin, 2009)
[09R1] U. Rössler, *Solid State Theory: An Introduction*, 2nd edn. (Springer, Berlin, 2009)

Chapter 15
Semiconductor Bandstructure

Abstract In order to understand the electronic and related optical properties of semiconductors one needs to analyze the details of their bandstructure. We will give a short review of some basic approaches to theoretically treat the bandstructure starting, e.g., from free electrons or from atomic orbitals, including the basic concepts of $\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory. With this understanding of the general properties of bandstructures we then take a look at some prominent substances. We will derive the detailed properties of the highest valence bands and lowest conduction band in GaAs, Si and other cubic semiconductors with tetrahedral coordination of the atoms/ions. This is followed by a discussion of wurtzite-type semiconductors like ZnO and GaN. We will close this chapter with some properties of further semiconductors, which have important applications in devices like narrow-gap semiconductors, oxides and new absorber materials for thin-film solar cell.

15.1 Basics of Bandstructure Calculations

In this section we will take a closer look at bandstructure calculations. But, we only want to present the basic concepts and ideas which are necessary to comprehend semiconductor optics. For a more stringent mathematical treatments of this topic we refer the reader to the literature and textbooks on solid-state and semiconductor physics like [72M1, 09R1, 16A1]. According to the discussion on electrons in periodic potentials in Sect. 14.2 we start with the concept of nearly free electrons and proceed to the approach starting from atomic orbitals. We will finish this section with the so-called $\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory, which is very useful to calculate bandstructure details in the vicinity of the extrema.

15.1.1 Nearly Free Electron Model

As already stated, the dispersion relation of a free electron in vacuum is a parabola in the non-relativistic regime as shown in Fig. 15.1a or Fig. 14.1a, i.e.,

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_15

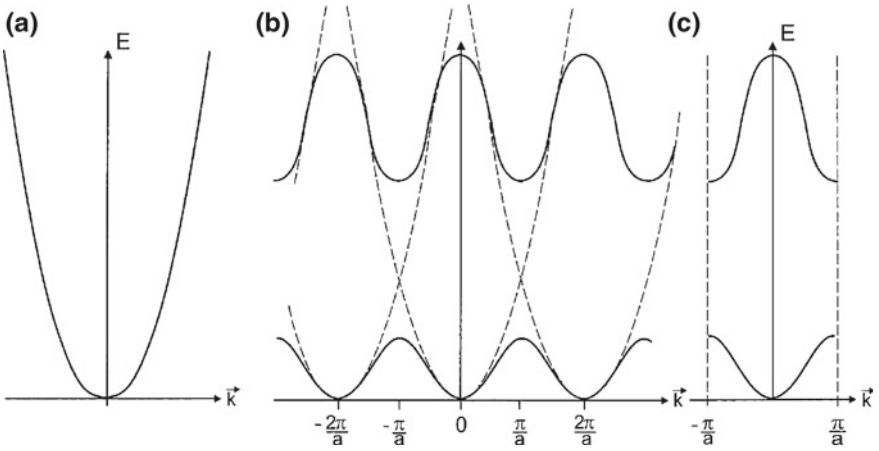


Fig. 15.1 The (one-dimensional) dispersion relation of a free electron in vacuum (**a**) in a weak periodic potential (**b**) and the resulting reduced-zone scheme (**c**)

$$E(\boldsymbol{k}) = \frac{\hbar^2 \boldsymbol{k}^2}{2m_0}. \tag{15.1}$$

We start, for simplicity, with a one-dimensional system and represent the relation (15.1) for this case in Fig. 15.1a.

Next we introduce a very weak periodic potential as shown schematically in Fig. 15.2 with a periodicity a. Weak means, in this context, that the dispersion relation of Fig. 15.1a is only marginally influenced by this potential. Nevertheless, we have introduced a periodicity by this potential in real space, which evokes the Bloch theorem. As a consequence, the dispersion relation can be repeated in $\boldsymbol{k}$-space by shifting it by integer multiples of $2\pi/a$ resulting in the extended zone scheme of Fig. 15.1b (dashed lines). Compare also with Fig. 14.1. In the limit that the amplitude of the periodic potential goes to zero, this is it. If the potential is small but finite, a further modification occurs. Small bandgaps open at the borders of the Brillouin zones as shown by the solid lines.

There are various possibilities to explain the appearance of these gaps and the fact that two different energy eigenvalues exist for the same $\boldsymbol{k}$-vector at the Brillouin zone boundary. A first argument reads that this splitting can be considered as an example of the non-crossing rule already discussed in Sect. 8.1. The coupling between the degenerate states at $k = \pm n\pi/a$ is mediated by the periodic potential. Indeed it can be shown that the width of the bandgap is determined by the Fourier component of the periodic potential (see, e.g., (9.5)) with the reciprocal lattice vector $\boldsymbol{G}$, which couples the bands that intersect for vanishing potential.

Another set of arguments works as follows. Assume a plane electron wave is trying to propagate in the periodic potential. Then the potential will produce scattered waves. For an arbitrary $\boldsymbol{k}$-vector, i.e., wavelength and (in dimensions higher than one) direc-

tion of incidence, the scattered waves will have various phase shifts relative to each other and will essentially cancel. However, there are special wave vectors, namely just the ones at the borders of the Brillouin zones (in our present one-dimension model $k = \pm n\pi/a$) where all scattered waves interfere constructively to form a back-scattered wave with large amplitude, i.e., a propagating wave with, e.g., $k_i = \pi/a$ creates a back-scattered wave with $k_\mathrm{b} = -\pi/a$ (and $k_i - k_\mathrm{b} = 2\pi/a = G$). Since the back-scattered wave with k_b scatters again off the periodic potential contributing to the wave with k_i the result under stationary conditions is two counter-propagating waves of equal amplitudes and $k_i = -k_\mathrm{b}$. As is well known, the superposition of two waves with $\boldsymbol{k}$ and $-\boldsymbol{k}$ and equal amplitudes produces a standing wave. This fact has in our context two consequences:

- A wave packet built up from a standing wave and its surroundings in $\boldsymbol{k}$-space has zero group velocity. Consequently the slope of the dispersion relation must be zero, normal to the border of the Brillouin zone.
- The standing wave can be the sine- or the cosine-type with the same wavelength and $\boldsymbol{k}$-vector as shown in Fig. 15.2, where the probabilities are plotted. These two waves have their maxima at the positions of the maxima or of the minima of the periodic potential, respectively. Consequently they see different average potential energies but have equal kinetic energy. This fact results in the energy splitting, i.e., in the appearance of the gaps at the borders of the Brillouin zones.

We proceed now, still in the context of nearly free electrons, to two- and three-dimensional periodic potentials.

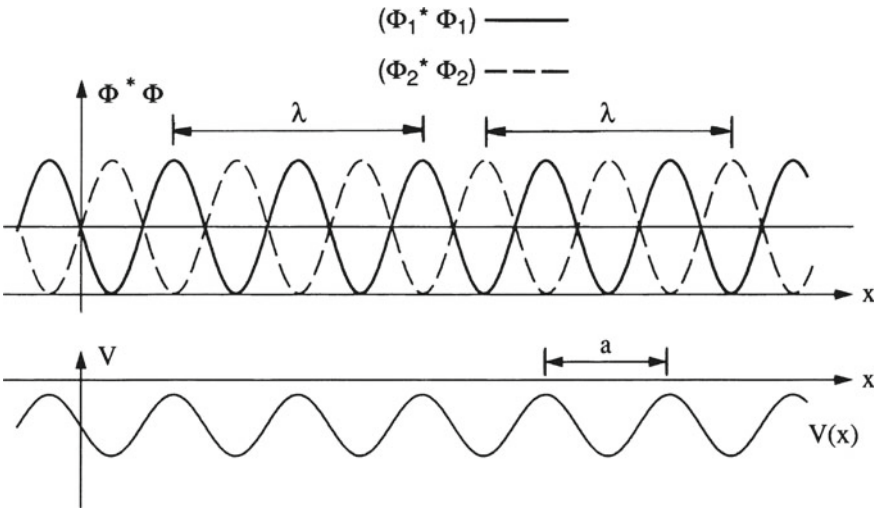


Fig. 15.2 Schematic drawing of two Bloch waves with equal $\boldsymbol{k}$-vector but different potential energies at the boundary of the first Brillouin zone over a periodic potential $V(x)$

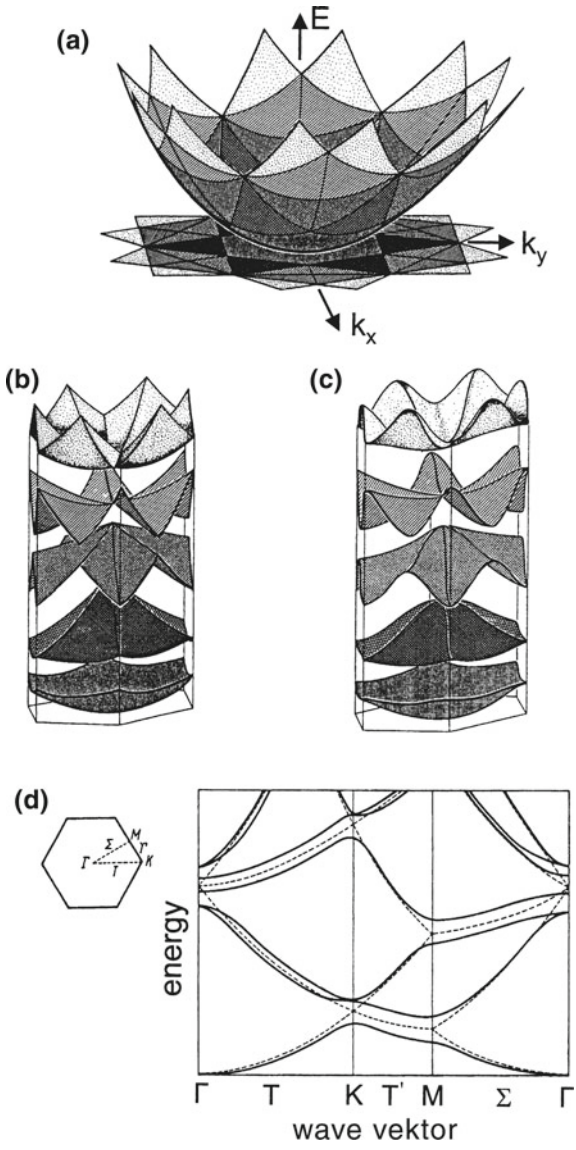


Fig. 15.3 The parabolic dispersion relation of free electrons over a two-dimensional $\boldsymbol{k}$-space, assuming a hexagonal lattice with vanishing potential (the so-called empty lattice) (**a**) the resulting reduced zone scheme for vanishing (**b**) and weak but finite potential (**c**) and the $E(k)$ relation along various lines of high symmetry for both cases (**d**). In **b** and **c** the bands are displaced vertically for clarity (reprinted with permission from [72M1] ©1972 Springer Nature)

In Fig. 15.3 we show the parabolic dispersion relation of (15.1) over a two-dimensional $\boldsymbol{k}$-space assuming a hexagonal, two-dimensional lattice but vanishing potential. The differently hatched areas in the k_x, k_y plane correspond to the different Brillouin zones. Figure 15.3a thus corresponds to the extended zone scheme. Figure 15.3b shows the resulting reduced zone scheme, which is obtained by shifting the parts outside the first Brillouin zone by suitable vectors $\boldsymbol{G}$ of the reciprocal lattice into the first zone. The various energy surfaces or bands touch, e.g., at the border of the

first Brillouin zone for vanishing potentials like in Fig. 15.1b for the one-dimensional case, but have been displaced in Fig. 15.3b, c for clarity along the energy axis to allow an "insight" into the first Brillouin zone.

In Fig. 15.3c a weak potential has been switched on. As we expect from our knowledge from the one-dimensional case the overall appearance of the $E(\boldsymbol{k})$ relation does not change very much but kinks are smoothed out, energy gaps open, e.g., at the boarder or the center of the first Brillouin zone and the slope of the dispersion relation vanishes at the zone boundary in the direction normal to it. Figure 15.3d finally gives the dispersion relation along various directions in the two-dimensional $\boldsymbol{k}$-space for vanishing and for weak potentials.

Empty-lattice bandstructure calculations for three-dimensional crystal structures (see dashed line in Fig. 15.3d already show a rather high degree of complexity). We notice that this complexity comes to a large extent from the back-folding of the parabolic dispersion into the first Brillouin-zone. Introduction of realistic potentials (as in Fig. 15.3d) and interactions between the electronic bands add more details to the bandstructure as we will see for real semiconductors in the next section.

Actually there is a possibility to test various bandstructure calculations, the so-called empty-lattice test. One runs any type of bandstructure calculation but with vanishing potential. Since the solution is known exactly, the bandstructure calculation must give this result. If it fails, something might be wrong, at least for the calculation of the bands arising from the outer, weakly bound atomic orbitals.

15.1.2 Tight-Binding Models

In the nearly free electron approach discussed so far we stressed the plane-wave factor of the Bloch-wave and assumed that the lattice periodic part is essentially a constant. We now start from the opposite side by placing one or more atomic orbitals at every atom site (*tight-binding approximation*) as shown, e.g., in Fig. 14.2 for simple 1s like atomic orbitals.

We translate this concept into a theoretical approach assuming, for simplicity, that we have a three-dimensional simple-cubic lattice with one atom per unit cell and a non-degenerate normalized atomic orbital $\varphi_i(\boldsymbol{r} - \boldsymbol{R}_i)$ at every lattice site $\boldsymbol{R}_i$. In the *linear combination of atomic orbitals (LCAO)*, the total wave function $\psi(\boldsymbol{r})$ is then given by

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) = \sum_i C_{\boldsymbol{k},i}\, \varphi_i(\boldsymbol{r} - \boldsymbol{R}_i)\,. \tag{15.2a}$$

The fact that this form has to be of the Bloch-type determines the $C_{\boldsymbol{k},i}$ to be

$$C_{\boldsymbol{k},i} = N^{-\frac{1}{2}} \mathrm{e}^{i\boldsymbol{k}\boldsymbol{R}_i} \tag{15.2b}$$

where $N^{-\frac{1}{2}}$ is the normalization factor for a crystal consisting of N atoms. Inserting (15.2b) in (15.2a) results in the LCAO wave function

$$|\boldsymbol{k}\rangle = \psi_{\boldsymbol{k}}(\boldsymbol{r}) = N^{-\frac{1}{2}} \sum_i \mathrm{e}^{\mathrm{i}\boldsymbol{k}\boldsymbol{R}_i} \varphi(\boldsymbol{r} - \boldsymbol{R}_i)\,. \tag{15.3a}$$

The energy associated to these wave functions can be calculated in first order with the help of the Hamiltonian H which contains, apart from the kinetic energy terms, the sum over the atomic potentials via

$$E(\boldsymbol{k}) = \langle \boldsymbol{k}\,|H|\,\boldsymbol{k}\rangle = N^{-1} \sum_i \sum_j \mathrm{e}^{\mathrm{i}\boldsymbol{k}(\boldsymbol{R}_i - \boldsymbol{R}_j)} \int \varphi^*(\boldsymbol{r} - \boldsymbol{R}_i) H \varphi(\boldsymbol{r} - \boldsymbol{R}_j) \mathrm{d}^3 r\,. \tag{15.3b}$$

Since only relative distances enter, we simplify (15.3b) by introducing $\boldsymbol{R}_m = \boldsymbol{R}_i - \boldsymbol{R}_j$ and obtain

$$E(\boldsymbol{k}) = \langle \boldsymbol{k}\,|H|\,\boldsymbol{k}\rangle = N^{-1} \sum_m \mathrm{e}^{\mathrm{i}\boldsymbol{k}\boldsymbol{R}_m} \int \varphi^*(\boldsymbol{r} - \boldsymbol{R}_m) H \varphi(\boldsymbol{r}) \mathrm{d}^3 r\,. \tag{15.3c}$$

The next simplification is to assume realistically that in the sum only the terms with $\boldsymbol{R}_m = 0$ and contribution for nearest neighbors $\boldsymbol{R}_n$ in the overlap integral give considerable contributions. With the abbreviations

$$\int \varphi^*(\boldsymbol{r}) H \varphi(\boldsymbol{r}) \mathrm{d}^3 r = -A \tag{15.4a}$$

and

$$\int \varphi^*(\boldsymbol{r} - R_n) H \varphi(\boldsymbol{r}) \mathrm{d}^3 r = -B \tag{15.4b}$$

we obtain

$$E(\boldsymbol{k}) = \langle \boldsymbol{k}\,|H|\,\boldsymbol{k}\rangle = -A - B \sum_n \mathrm{e}^{\mathrm{i}\boldsymbol{k}\boldsymbol{R}_n} \tag{15.5}$$

where the sum now runs only over the nearest neighbors.

The term $-A$ is close to but generally slightly lower in energy than the eigenenergy of the parent atomic orbital $\varphi(\boldsymbol{r})$ since the Hamiltonian contains not only the potential of a single atom but the sum over them as stated above.

For the simple-cubic lattice, which we consider, $\boldsymbol{R}_n$ adapts the following six values

$$\boldsymbol{R}_n = (\pm a, 0, 0);\ (0, \pm a, 0);\ (0, 0, \pm a)\,. \tag{15.6}$$

This finally results in a simple bandstructure of width $12B$ centered around $-A$:

$$E(\boldsymbol{k}) = -A - 2B(\cos k_x a + \cos k_y a + \cos k_2 a)\,. \tag{15.7}$$

We see from (15.7) that the overlap between atomic orbitals leads to the formation of bands out of sharp atomic levels while the nearly free electron methods explained the appearance of gaps in the simple parabolic dispersion relation (see again Fig. 14.1). Physically the formation of bands by the overlap described by B in (15.7) is nothing but the splitting of the eigenenergies of coupled harmonic oscillators as already mentioned in Sect. 14.1.

Again (as in Sect. 14.5) we can argue that the band width increases with increasing overlap integral B. Consequently, the width of the forbidden gaps tends to decrease, as does the effective mass, and the lattice-periodic part of the Bloch-function $u_k(\boldsymbol{r})$ is smoothed out.

An early example of the LCAO method is the *Wigner–Seitz approach* [47L1] which demonstrated for the simple alkali-metals, that the nearly free electron and the LCAO approaches give consistent results.

To finish this introduction into LCAO we want to discuss two further examples. First we consider qualitatively a hypothetic situation in which a valence-band is formed from atomic $2p$ orbitals in a simple-cubic lattice.

In Fig. 15.4a we show a layer of p_x orbitals. In the x direction the overlap integral B is small since adjacent wave function lobes of different atoms have opposite sign, i.e., there is a node line. Small overlap means a narrow band in the k_x direction and a heavy hole mass. The same orbitals have a large overlap integral in the y direction since adjacent lobes of neighboring atoms have the same sign. Consequently, B is larger in the y direction, the band is broader in the k_y direction, and the effective mass is small as shown by the solid line in Fig. 15.4b. For the p_y orbitals we obtain just the opposite behavior as shown by the dashed line in Fig. 15.4b. Including the p_z orbitals gives an analogous behavior in all three directions of $\boldsymbol{k}$-space. We can reinterpret this

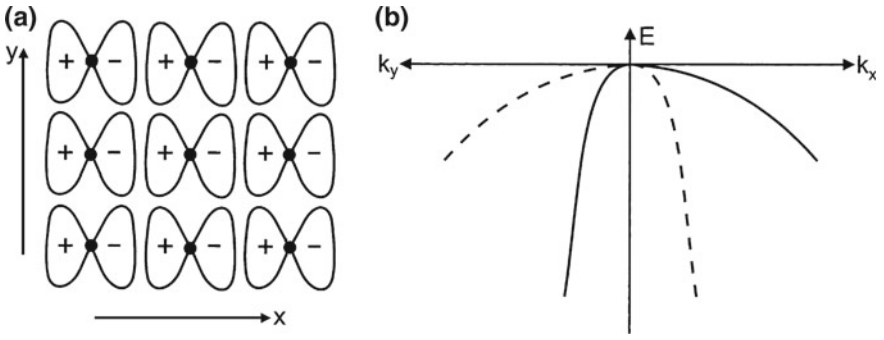


Fig. 15.4 A hypothetic layer of p_x orbitals (**a**) and the resulting $E(\boldsymbol{k})$ relation in the direction k_x and k_y (*solid line*) and the analogue for p_y orbitals (*dashed line*) (**b**). The *heavy dots* in **a** give the positions of the atoms

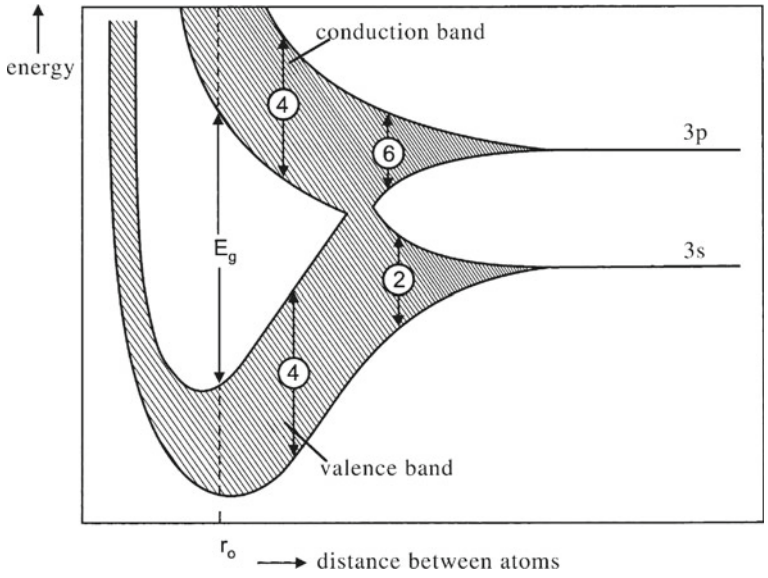


Fig. 15.5 The bands originating from the $3s$ and $3p$ levels of Si as a function of the lattice constant a or the distance between atoms. The numbers in the circles give the number of electrons per atom in the respective state or band (adapted with permission from [09I1] ©2009 Springer Nature)

finding in the following way. There is one valence-band in all three directions, which has low curvature, i.e., a heavy effective mass and another one, which is degenerate at the Γ-point, with large curvature and low effective mass. Consequently the holes in these bands are called *heavy holes* (hh) and *light holes* (lh).

We should remember that we have been treating a hypothetical material, but we shall encounter in the next chapter a very similar but slightly more complex situation for the valence bands of many semiconductors crystallizing in diamond- or zincblende-type structures.

As second example we discuss with Fig. 15.5 the tight-binding approach for Si (or diamond or Ge) as a function of the lattice constant a. Si has a filled Ne configuration plus $3s^2 3p^2$, i.e., four electrons in the outer $n_B = 3$ shell, which are responsible for the chemical binding. The lower electron states ($1s^2$, $2s^2$, $2p^6$) form deeper valence bands, which are presently of no interest.

For a large lattice constant the overlap integral vanishes and one starts with the atomic $3s$ and $3p$ levels. With decreasing distance the levels shift to lower energy and broaden since both A and B in (15.7) depend on the lattice constant. For further decreasing distance the two bands start to overlap. This is the distance between the atoms at which the sp^3 hybrid orbitals start to form. For even smaller values the bands split again. In contrast to the twofold degenerate $3s$ and sixfold degenerate $3p$ levels, two fourfold degenerate bands form the binding and the antibinding states of the sp^3 hybrids. Since there are two Si-atoms in the primitive unit cell, the valence band

accommodates the eight outer electrons while the conduction band also has place for eight electrons per unit cell, but stays unoccupied (at $T = 0\,\mathrm{K}$). The binding orbitals of the sp^3 hybrid point towards the corners of a tetrahedron giving rise to the tetrahedral coordination of many semiconductors (see Sect. 9.1).

The minimum of the total free energy gives the equilibrium distance r_0, which is indicated in Fig. 15.5. For even smaller values of a the bands shift to higher energies, again due to the action of Pauli's principle, and the increasing kinetic energy of the wave functions.

15.1.3 $k \cdot p$ Perturbation Theory

We now want to give a few introductory remarks on the widely used $\boldsymbol{k} \cdot \boldsymbol{p}$ *perturbation theory*, which is very valuable to calculate the bandstructure in the vicinity of band extrema. Also by extrapolation it allows one obtaining an idea about the bandstructure over the whole Brillouin zone. For more details see again the textbooks on solid state physics and on semiconductor physics cited in this chapter or [57K1, 66K1, 10Y1].

We assume in the following that the band extrema are situated at $\boldsymbol{k}_0$ and introduce the further simplification that $\boldsymbol{k}_0 = 0$. If this is not the case, the following considerations are valid for $\boldsymbol{k}' = \boldsymbol{k} - \boldsymbol{k}_0$. Furthermore we neglect spin.

We start again with the one-electron approximation and the periodic potential $V(\boldsymbol{r}) = V(\boldsymbol{r} + \boldsymbol{R})$. The Hamiltonian reads

$$H = \frac{\boldsymbol{p}^2}{2m_0} + V(\boldsymbol{r}). \tag{15.8}$$

We insert the Bloch-wave ansatz of Sect. 14.2 into (15.8) and obtain (15.9a) where n is the band index, m_0 the free electron mass and $\boldsymbol{p} = \frac{\hbar}{i}\nabla$:

$$\left(\frac{\boldsymbol{p}^2}{2m_0} + \frac{\hbar \boldsymbol{k} \cdot \boldsymbol{p}}{m_0} + \frac{\hbar^2 \boldsymbol{k}^2}{2m_0} + V(\boldsymbol{r})\right) u_{n,\boldsymbol{k}}(\boldsymbol{r}) = E_n(\boldsymbol{k}) u_{n,\boldsymbol{k}}(\boldsymbol{r})\,. \tag{15.9a}$$

For the case $\boldsymbol{k} = \boldsymbol{k}_0 = 0$ (15.9a) simplifies to

$$\left(\frac{\boldsymbol{p}^2}{2m_0} + V(\boldsymbol{r})\right) u_{n,0}(\boldsymbol{r}) = E_n(\boldsymbol{k} = 0) u_{n,0}(\boldsymbol{r}) = E_{n,0} u_{n,0}(\boldsymbol{r})\,. \tag{15.9b}$$

We now assume, that the solution of (15.9b) is known. Then we can use the terms $\hbar \boldsymbol{k} \cdot \boldsymbol{p}/m_0$ (this term is the reason for the name of this method) and $\hbar^2 \boldsymbol{k}^2/2m_0$ in (15.9a) as perturbations of the first and second order in $\boldsymbol{k}$.

For the simple case of a non-degenerate band (or a twofold spin-degenerate one) we obtain for the lattice periodic part of the eigenfunctions $u_{n,\boldsymbol{k}}(\boldsymbol{r})$

$$u_{n,\boldsymbol{k}}(\boldsymbol{r}) = u_{n,0}(\boldsymbol{r}) + \frac{\hbar}{m_0} \sum_{n' \neq n} \frac{\langle u_{n,o} \left| \boldsymbol{k} \cdot \boldsymbol{p} \right| u_{n',0} \rangle}{E_{n,0} - E_{n',0}} u_{n',0}(\boldsymbol{r}) \,. \tag{15.10a}$$

The above-mentioned situation of non-degeneracy frequently holds for the lowest conduction band of semiconductors. The dispersion relation is then given by

$$E_n(\boldsymbol{k}) = E_{n,0} + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar^2}{m_0^2} \sum_{n' \neq n} \frac{\left| \langle u_{n,0} \left| \boldsymbol{k} \cdot \boldsymbol{p} \right| u_{n',0} \rangle \right|^2}{E_{n,0} - E_{n',0}} \,. \tag{15.10b}$$

From (15.10b) we can easily deduce the effective mass of, e.g., the conduction band n

$$\frac{1}{m_\mathrm{e}} = \frac{1}{m_0} + \frac{2}{m_0^2 \boldsymbol{k}^2} \sum_{n' \neq n} \frac{\left| \langle u_{n,0} \left| \boldsymbol{k} \cdot \boldsymbol{p} \right| u_{n',0} \rangle \right|^2}{E_{n,0} - E_{n',0}} \,. \tag{15.10c}$$

We notice that the $\boldsymbol{k} \cdot \boldsymbol{p}$-term leads to a coupling between various bands. This coupling influences the effective electron mass m_e. The momentum matrix element in turn is closely related to the dipole-matrix element for the transition between bands n and n' (see Sect. 6.2) and can at least in principle be deduced from optical absorption spectra.

The knowledge, which of the matrix elements vanish (and thus do not lead to band interactions) and which are non-zero, can be obtained from group theory (see Appendix B). The coupling strength scales with the inverse of the energetic separation $E_{n,0} - E_{n',0}$ of the respective bands. If the band indices n and n' in (15.10c) stand for the lowest conduction and highest valence band, respectively, we find the reduction of the effective electron mass with decreasing width of the gap $E_{n,0} - E_{n',0}$—as already stated various times above.

Since the $\boldsymbol{k} \cdot \boldsymbol{p}$-theory starts for all bands with the free electron mass (see (15.9a)) the appearance of effective masses can be considered again as a consequence of level repulsion, e.g., between conduction and valence band in the way indicated schematically by the arrows in Fig. 15.6. Even the change of sign of the valence-band curvature results from the band interaction.

The number of bands at $\boldsymbol{k}_0$ is infinite, but due to the energy denominator in (15.10) typically only few close-lying bands need to be considered. The treatment of degenerate bands, especially of the upper valence bands in $\boldsymbol{k} \cdot \boldsymbol{p}$ is more complicated. Also the electron spin needs to be considered. For semiconductors with sp^3-hybridized binding often the eight-band Kane model is applied see, e.g., [57K1, 84R1, 09R1]. We will discuss these aspects in some more detail for the bandstructure of real semiconductors in the following section.

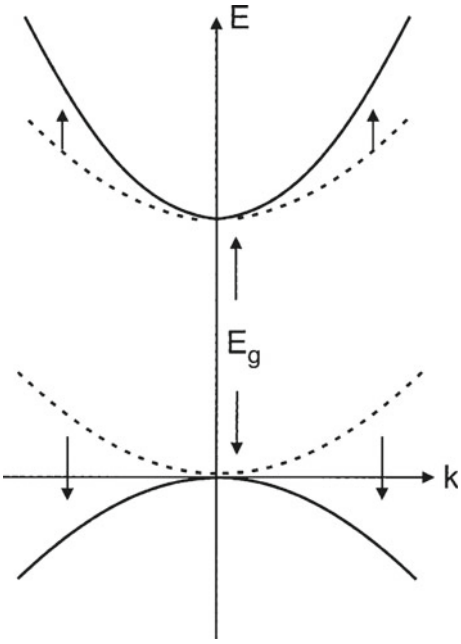


Fig. 15.6 Conduction and valence band with free electron mass m_0 (*dashed lines*) and the modifications (*arrows*) resulting from level repulsion (*solid lines*)

15.2 Bandstructure of Real Semiconductors

After we have elaborated general properties of bandstructures we want to describe now the bandstructures of real semiconductors. We will chose here as examples the most important semiconductors for opto-electronic applications as well as some prominent substances being characteristic for certain crystalline structure and/or binding.

15.2.1 Semiconductors with Diamond or Zincblende Structure

We start with (dominantly) covalent-bound semiconductors of diamond structure (elemental semiconductors Si or Ge; point group O_h) and of zincblende structure for III–V, II–VI and I–VII compounds (e.g., compound semiconductors GaAs, CdTe or ZnSe; point group T_d). Due to their similarity in crystal structure and binding the bandstructures of these semiconductors have some common features. Most importantly the upper valence bands arise essentially from the bonding states of the p-orbitals of the anions (or atoms in the case of Si and Ge). The lowest conduction bands come from the lowest empty s-levels of the cations/atoms or (in the case of Si) the anti-bonding p-orbitals, respectively.

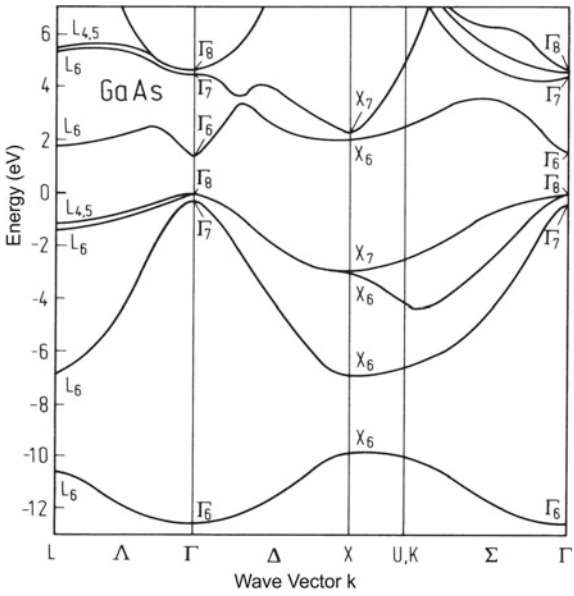


Fig. 15.7 Bandstructure of GaAs including spin-orbit coupling (adapted with permission from [88C1] ©1988 Springer Nature, see also [10Y1])

Calculated bandstructures of GaAs and Si are depicted in Figs. 15.7 and 15.8, respectively. The bandstructures are calculated for the highest valence and lowest conduction bands along lines in reciprocal space between points of high symmetry. For the notation of lines and points in the first Brillouin zone see Fig. 9.4. The labeling of the minima and maxima of the bands follows the notation of group theory, which is explained in detail in Appendix D. The notations for the bandstructure features differ in these two figures because spin-orbit interaction is neglected in the calculations for Si (Fig. 15.8). Consideration of spin leads to the so-called double-group notation in Fig. 15.7. The bandstructures shown in Figs. 15.7 and 15.8 have been calculated using pseudo-potentials (see Sect. 14.2). For a detailed calculation of these bandstructures using a thirty-band $\boldsymbol{k}\cdot\boldsymbol{p}$-model see [04R1].

The bandstructure of diamond and zincblende semiconductors like GaAs, InAs, Si, and Ge have many common features as we will illustrate with Fig. 15.9. We will now focus exemplary on details of the bandstructure of GaAs (Fig. 15.7).

The valence band of the diamond and zincblende semiconductors has its maximum at the Γ point, i.e., at $\boldsymbol{k} = 0$. It is sixfold degenerate including spin corresponding to the parent p orbitals. In Sect. 15.1 we stated that eight levels form the bonding sp^3 hybrid state since there are two atoms in a unit cell. Six of them, which are treated in the following, form the top of the valence band and the other two the rather low lying Γ_6 level seen, e.g., in Fig. 15.7. This latter band is not important for semiconducting properties and thus not considered further in the following. The upper valence band splits due to spin-orbit coupling at $\boldsymbol{k} = 0$ into a twofold degenerate band (symmetry Γ_7^+ in O_h or Γ_7 in T_d) and a fourfold degenerate band (symmetry Γ_8^+ in O_h or Γ_8 in T_d). Usually the Γ_8 band is the upper and the Γ_7 the lower one. The spin-orbit

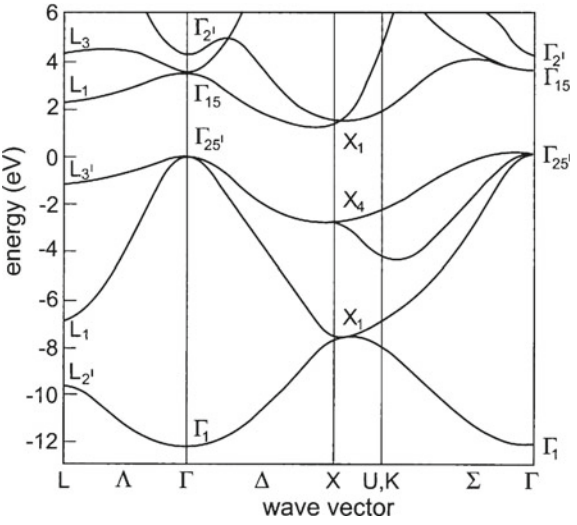


Fig. 15.8 Bandstucture of Si (adapted with permission from [88C1] ©1988 Springer Nature, see also [10Y1])

splitting increases in atoms with increasing charge Z of the nucleus and this also applies to semiconductors. This is the reason why it is small in Si [04R1] (it actually was neglected in Fig. 15.8) in comparison to Ge or GaAs.

The Γ^8 valence band splits for $\boldsymbol{k} \neq 0$ into two bands which have different curvature and are therefore known as *heavy-hole* and *light-hole* bands (see the discussion related to Fig. 15.4. All bands have cubic symmetry, which is lower than spherical symmetry. As a consequence the dispersion and thus the hole masses depend on the direction of $\boldsymbol{k}$. This phenomenon is known as *band-warping*.

The valence bands are frequently treated in $\boldsymbol{k} \cdot \boldsymbol{p}$ perturbation theory (see [66K1, 04R1, 09R1, 10Y1] and Sect. 15.1.3). Basically one has to start with the 16 bonding and anti-bonding states resulting from two atoms per primitive unit cell. Often one considers the lowest twofold degenerate conduction band (symmetry Γ_6) separately, e.g., in the way described in Sect. 15.1.3 and is then left with eight valence-band states. Since the lowest Γ_6 state is energetically rather far away, it has only minor influence and may be neglected in the treatment of the upper valence bands. The remaining six states in the Γ_8 and Γ_7 valence bands can be treated in a 6×6 matrix. If the spin-orbit coupling is large, it is, however, often possible to treat the Γ_8 band alone resulting in a 4×4 matrix, which can be solved analytically. The resulting bandstructure is frequently described by the so-called *Luttinger parameters* γ_1, γ_2 and and γ_3 according to

$$E_{1,2} = E_0 + \frac{\hbar^2}{2m_0}\Big[\gamma_1 k^2 \pm 2[\gamma_2(k_x^4 + k_y^4 + k_z^4) + 3\left(\gamma_3^2 - \gamma_2^2\right)\left(k_x^2 k_y^2 + k_y^2 k_z^2 + k_z^2 k_x^2\right)]^{1/2}\Big] \tag{15.11a}$$

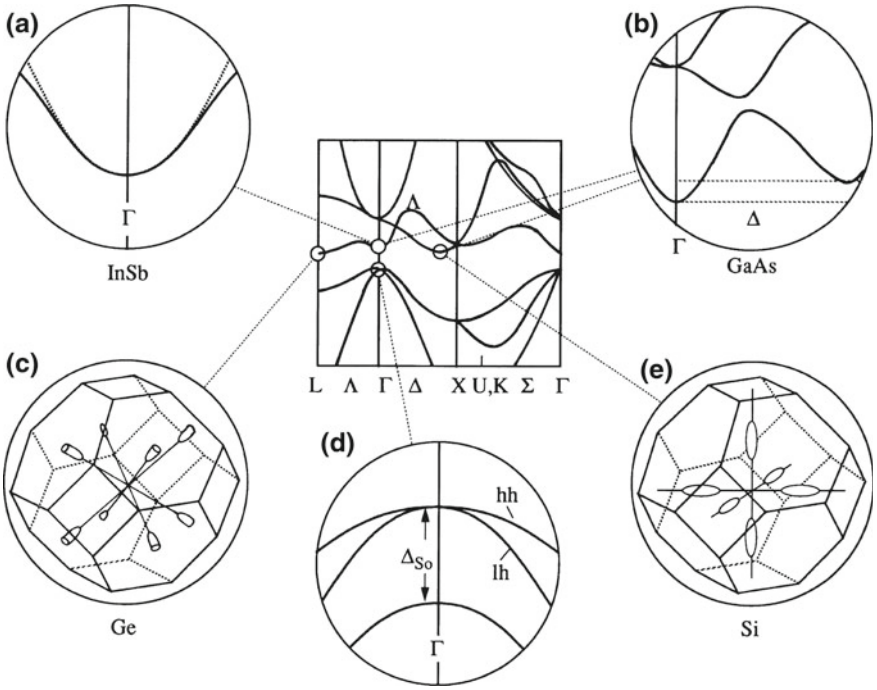


Fig. 15.9 An overview of the bandstructures of some cubic semiconductors demonstrating non-parabolicity of the conduction band (**a**), avoided crossing of bands (**b**), surfaces of constant energy at the L point in Ge (**c**), the highest valence bands at the Γ-point (**d**), and surfaces of constant energy close to the X point in Si (**e**) (adapted with permission from [70M1] ©1970 Springer Nature)

where γ_1^{-1} describes the average effective mass and γ_2 and γ_3 the splitting into heavy- and light-hole bands and the warping via the term $3\left(\gamma_3^2 - \gamma_2^2\right)$. If warping is negligible (i.e. $\gamma_2 = \gamma_3$) the light-hole and heavy-hole effective masses are given by

$$m_{\mathrm{hh}} = \frac{m_0}{\gamma_1 - 2\gamma_2}, \quad m_{\mathrm{lh}} = \frac{m_0}{\gamma_1 + 2\gamma_2} \tag{15.11b}$$

Also in the case of zincblende structure but dominant ionic binding, the valence bands stem from the highest occupied atomic p-levels of the anions with a more-or-less pronounced admixture of d-levels. In CuCl the ordering of the valence bands is even inverted due to the influence of close-lying d-levels [63C1, 65S1].

Let us now consider the conduction band of cubic (diamond or zincblende structure) semiconductors. The lowest conduction band typically has a minimum at the Γ point (with the exception of Si) and further minima in the direction Δ close to the X points and at the L points. If the minimum of the Γ point is the deepest one, the semiconductor is said to have a *direct gap* as mentioned above, since transitions between the global maximum of the valence band and the global minimum of the

conduction band are directly possible with photons, having $k_{\text{photon}} \simeq 0$. The other cases are called *indirect-gap* semiconductors since a mechanisms is required to fulfill momentum conservation in optical transitions. Often this achieved by involving an additional phonon (like in Si and AlAs). But the momentum conservation can also be relaxed, e.g., by disorder in alloys like $Al_{1-y}Ga_yAs$ leading to zero-phonon transitions across the indirect gap.

Examples of indirect semiconductors are diamond, Si and Ge, some of the III–V compounds such as AlAs and GaP, and some of the I–VII compounds like AgBr. Direct-gap semiconductors include some of the III–V compounds like GaAs, InAs or InP, the II–VI compounds ZnS, ZnSe, ZnTe and CdTe, and I–VII materials like CuCl, CuBr and CuI. See also Table 1.1.

The conduction band minimum at the Γ point is usually, to a very good approximation, isotropic and parabolic. Only some narrow-gap materials like InSb show significant non-parabolicities in the vicinity of $k = 0$ (Fig. 15.9).

The minima at the L points of Ge or along the Δ direction of Si (Fig. 15.8) are parabolic but highly anisotropic. Surfaces of constant energy around these extrema are displayed in (Fig. 15.9). The dispersion relation around the respective minimum at $\boldsymbol{k}_0$ can be expressed by

$$E(\boldsymbol{k}) = E_{\text{g}} + \frac{\hbar^2}{2}\left(\frac{(k_x - k_{0x})^2}{m_{\text{l}}} + \frac{(k_y - k_{0y})^2}{m_{\text{t}}} + \frac{(k_z - k_{0z})^2}{m_{\text{t}}}\right). \qquad (15.12)$$

Here m_{l} is the *longitudinal effective mass* for k-components in the direction from Γ to $\boldsymbol{k}_0$ and m_{t} that for the two directions perpendicular to it (*transverse effective mass*).

Direct gap semiconductors are often called *single-valley semiconductors*, and the indirect ones *multi-valley semiconductors* because they have several (6 for Si and $8 \cdot 1/2$ for Ge) equivalent conduction-band minima (see Fig. 15.9).

But, multi-valley situations can also arise for other other scenarios [96K1]. In GaAs the minimum at the Γ point is the deepest, but other minima at different points of the Brillouin zone are close in energy, e.g., in the Δ direction. This gives rise to a negative differential resistance (utilized in the Gunn effect), which arises from the transfer of electrons under the influence of a strong electric field from the minimum at $k = 0$ with low effective mass to side minima with higher effective masses.

Another example is the alloy semiconductor $Al_xGa_{1-x}As$ which shows a transition from a direct-gap to an indirect-gap semiconductor at $x \approx 0.42$. The scenario around this crossover features significant coupling between the valleys at different points of the Brillouin zone leading to efficient inter-valley scattering of electrons mediated by phonons and disorder or even the occurrence of efficient stimulated emission across an indirect bandgap [96K1].

Some typical features of cubic diamond and zincblende semiconductors are summarized in Fig. 15.9.

15.2.2 Semiconductors with Wurtzite Structure

Semiconductors with hexagonal wurtzite structure (point group C_{6v}) are usually direct-gap semiconductors. They are found preferentially among II–VI compounds such as ZnO, ZnS, CdS and CdSe, but also among III–V materials like GaN.

The upper six valence-band states of the C_{6v} semiconductors are split by spin-orbit coupling and by the hexagonal crystal field into three subbands which are usually labeled from higher to lower energies as A, B and C bands symmetries Γ_9, Γ_7 and Γ_7 (Fig. 15.10a). In ZnO, the symmetries of the two upper bands are most probably inverted as in the case of CuCl (for discussions and literature see, e.g., [69R1, 10K1, 10K2]). The effective masses of the valence bands are often strongly anisotropic in these compounds, $m_\perp$ usually being smaller than $m_\parallel$. Here the indices refer to $\boldsymbol{k}$-vectors perpendicular and parallel to the polar crystallographic $\boldsymbol{c}$-axis. The effective mass entering the density of states has then to be calculated (similar to the case of longitudinal and transverse mass in Si and Ge conduction-band minima) by:

$$m_{\mathrm{DOS}} = (m_\perp^2 m_\parallel)^{1/3}. \tag{15.13}$$

If one restricts the discussion to the four states in the A and B bands, $\boldsymbol{k}\cdot\boldsymbol{p}$ theory predicts for negligible warping, i.e., $\gamma_2 = \gamma_3 = \gamma$ for the effective masses [92E1]

$$m_{\mathrm{A}\parallel} = \frac{m_0}{\gamma_1 - 2\gamma}; \quad m_{\mathrm{A}\perp} = \frac{m_0}{\gamma_1 + \gamma} \tag{15.14a}$$

$$m_{\mathrm{B}\parallel} = \frac{m_0}{\gamma_1 + 2\gamma}; \quad m_{\mathrm{B}\perp} = \frac{m_0}{\gamma_1 - \gamma}. \tag{15.14b}$$

A peculiarity of the states of Γ_7 symmetry is that they can have a *term linear in k* in the dispersion relation for $\boldsymbol{k}\perp\boldsymbol{c}$ as shown in Fig. 15.10b. This term has some influence on the optical properties. In principle, it can also occur for the Γ_7 conduction band, but is much smaller there and usually neglected. The same is true for the hexagonal warping which is, in principle, possible for C_{6v} symmetry in the plane normal to $\boldsymbol{c}$.

As already mentioned, the crystal structures of zincblende and of wurtzite are rather similar and differ only in the arrangements of the next-nearest neighbors. The unit cell of the wurtzite structure is in one direction twice as long as the primitive unit cell of the zincblende structure. As a consequence the first Brillouin zone is only half as long in one direction. The resulting folding back of the dispersion is shown schematically in Fig. 15.11 neglecting spin [69R1] for the two lowest conduction bands and the highest valence band. Obviously the number of bands doubles in this procedure, as does the number of states at the Γ point. This is in agreement with the fact that diamond and zincblende structures have two atoms in the basis of the primitive unit cell but wurtzite has four.

Due to the p- and s-type character of valence and conduction bands, for the more ionic bound semiconductors and of the binding and anti-binding sp^3 states close to

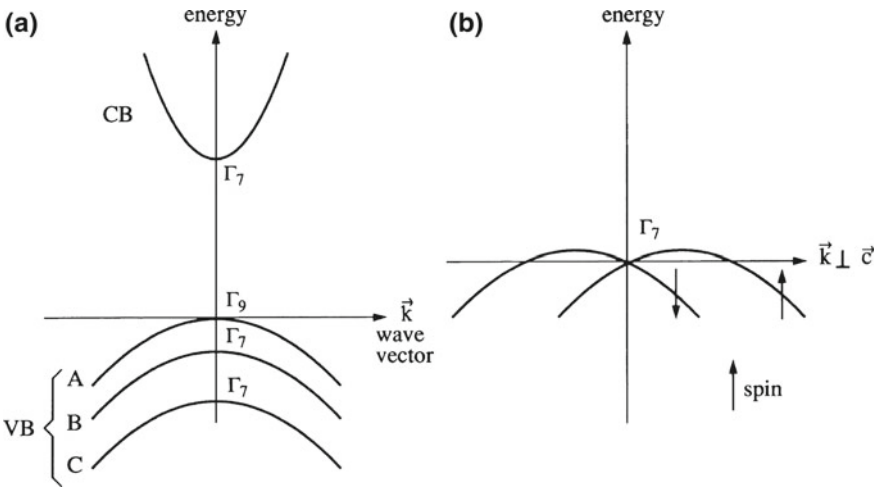


Fig. 15.10 Details of the bandstructure of hexagonal semiconductors around the Γ point. The splitting into three valence bands (**a**) and the influence of a term linear in k (**b**)

the fundamental gap, respectively, the band-to-band transition is dipole allowed—possibly with some additional selection rules for the hexagonal symmetry.

15.2.3 Further Semiconductors Important for Applications

We finally want to highlight some semiconductors not introduced so far. We will restrict this discussion to a couple of materials important for opto-electronic applications. They all show properties that deviate from the ones described above.

First of all there are the lead salts (PbS, PbSe, PbTe) which crystallize in the NaCl structure. They are narrow-gap semiconductors and have a direct gap. But, in contrast to the above-mentioned direct semiconductors, the band extrema are situated at the L points.

In Fig. 15.12 we illustrate the properties of the zincblende alloy semiconductor $Cd_{1-y}Hg_yTe$. It features a transition from a normal semiconductor ($y = 0$) to a narrow-gap material for $0 < y < 0.85$, and finally to a semi-metal for $1 \geq y \geq 0.85$.

Some semiconductors, such as the IV–VI compounds TiO_2 (rutile) and SnO_2 or the I–VI compound Cu_2O also have a direct gap. The transitions between the band extrema are, however, dipole forbidden because the extrema have equal parity. The symmetry groups of these materials contain inversion as an element and parity is therefore a good quantum number. This feature has significant consequences for the optical properties which we will discuss detail in Appendix B.

Indium tin oxide (ITO) is an important transparent electrode material used in opto-electronic devices. Doping of Sn into In_2O_3 leads to high electron conductivity.

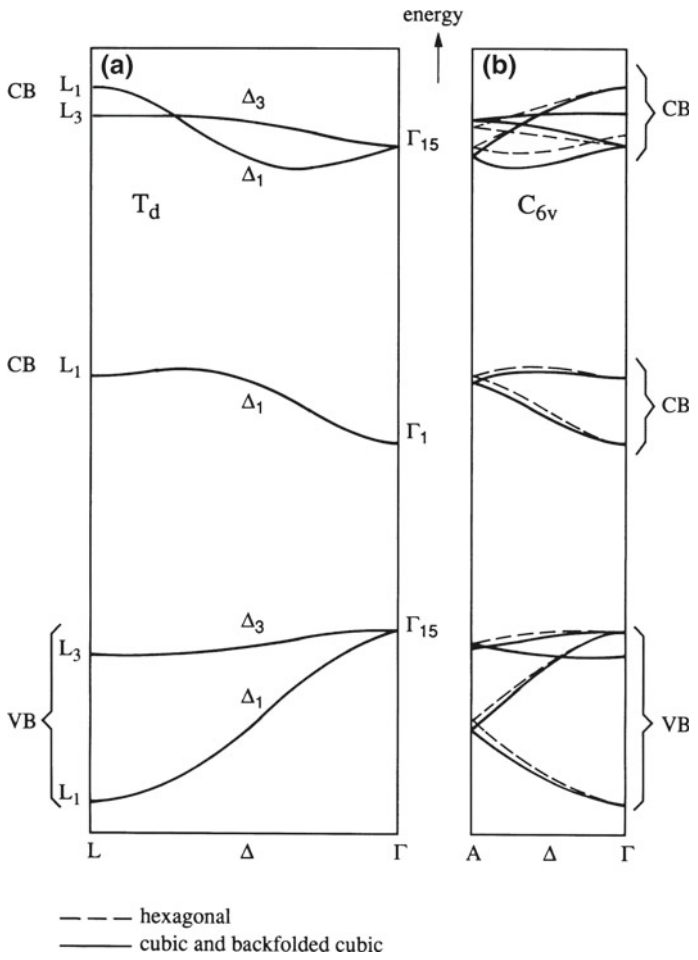


Fig. 15.11 The relation between the bandstructure of semiconductors with zincblende and wurtzite type structures along the corresponding direction in $\boldsymbol{k}$-space, neglecting spin (reprinted with permission from [69R1] ©1969 American Physical Society https://doi.org/10.1103/PhysRev.184.733)

Some experimental studies indicate an indirect nature of the bandgap of this material. A valence-band maximum off the Γ-point can result from mixing of O 2p and In 4d orbitals. But experimentally determined gaps might be determined by defect-related transitions. For a review and discussion see [07E1].

We want to close with two material systems which have applications as absorber materials in thin-film solar cells. The first one is already well established: Cu(In,Ga)Se_2 (CIGS) [11S1]. CIGS crystallizes in the tetragonal chalcopyrite structure. The latter is essentially a superstructure of two zincblende unit cells which still has a tetrahedral arrangement of the atoms. This close resemblance of the zincblende

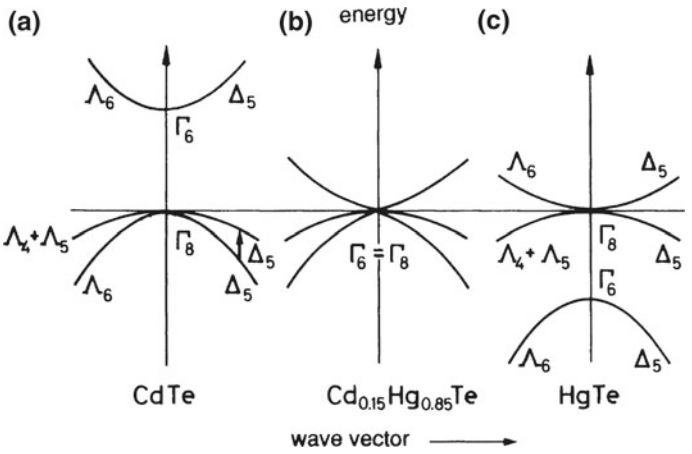


Fig. 15.12 The transition from a semiconductor (**a**) to a semi-metal (**c**) for the alloy Cd_{1-y} Hg_y Te (**b**) as a function of the composition y (adapted with permission from [70M1] ©1970 Springer Nature)

structure results in similar bandstructures which is, however, smaller than typical zincblende II–VI semiconductors. This is due to a slight tetragonal distortion of the lattice and the nature of the valence band arising from the hybridization of Cu-3d and Se-4p orbitals. The valence band is split into three bands due to the crystal field and the spin-orbit coupling [73S1, 83J1].

Hybrid organic/inorganic lead halides are a new class of possible absorber materials that have lately been studied intensively [14G1]. An example of this class is $CH_3NH_3PbX_3$ (with X = Cl, I, Br) incorporating large organic methylamonium cations. These materials crystallize in the polar perovskite structure, where the cations are surrounded by eight halide (anion) octaeders encasing each a small lead cation. Strong excitonic effects are found at the the direct bandgap, which however, has an inverted ordering of the bands with a split p-type conduction band [12E1].

Problems

15.1. Find out about the electron configuration of valence electrons in the atoms Si, Ge, Ga, In, Al, Zn, Cd, Cu, As, Se, S, O, N, Cl … and determine the orbitals forming the Γ-point valence and conduction bands in the related semiconductors (Si, Ge, GaAs, GaN, ZnSe, Cu_2O, …).

15.2. Make a sketch of surfaces of constant energy in a two- (or even three-) dimensional $\boldsymbol{k}$-space for spherical, simple cubic, and hexagonal (plane $\boldsymbol{k} \perp \boldsymbol{c}$) symmetries. Is spherical symmetry compatible with cubic and hexagonal symmetry?

15.3. Show that the wave vectors for which the scattered waves interfere constructively in one- and two-dimensional square lattices in the concept of nearly free electrons are just the borders of the Brillouin zones.

15.4. Verify that (15.2a) fulfils the Bloch criterion, i.e., that the wave function is transformed onto itself under a translation $\boldsymbol{R}_i$ apart from a phase factor.

References

[47L1] F.C. von der Lage, H.A. Bethe, Phys. Rev. **71**, 612 (1947)
[57K1] E.O. Kane, J. Phys. Chem. Solid **1**, 249 (1957)
[63C1] M. Cardona, J. Phys. Chem. Solid **24**, 1543 (1963); Phys. Rev. **129**, 69 (1963)
[65S1] K. Shindo, A. Morita, H. Kamimura, J. Phys. Soc. Jpn. **20**, 2054 (1965)
[66K1] E.O. Kane, The $k \cdot p$ method, in *Physics of III-V Compounds, Volume 1 of Semiconductors and Semimetals*, ed. by R.K. Willardson, A.C. Beer (Academic, New York, 1966), pp. 75–100
[69R1] U. Rössler, Phys. Rev. **184**, 733 (1969)
[70M1] O. Madelung, *Grundlagen der Halbleiterphysik*. Heidelberger Taschenbücher, vol. 71 (Springer, Berlin, 1970)
[72M1] O. Madelung, *Festkörpertheorie I: Elementare Anregungen* (Springer, Berlin, 1972)
[73S1] J.L. Shay, B. Tell, Surf. Sci. **37**, 748 (1973)
[83J1] J. Jaffe, A. Zunger, Phys. Rev. B **28**, 5822 (1983)
[84R1] U. Rössler, Solid State Commun. **49**, 943 (1984)
[88C1] M. Cohen, J.R. Chelikowsky, *Electronic Structure and Optical Properties of Semiconductors*. Springer Series in Solid-State Sciences, vol. 75 (Springer, Berlin, 1988)
[92E1] A.L. Efros, Phys. Rev. B **46**, 7448 (1992)
[96K1] H. Kalt, *Optical Properties of III–V Semiconductors: The Influence of Multi-valley Band Structures*. Springer Series in Solid-State Sciences, vol. 120 (Springer, Berlin, 1996)
[04R1] S. Richard, F. Aniel, G. Fishman, Phys. Rev. B **70**, 1–6 (2004)
[07E1] P. Erhart et al., Phys. Rev. B **75**, 153205 (2007)
[09I1] H. Ibach, H. Lüth, *Solid State Physics* (Springer, Berlin, 2009)
[09R1] U. Rössler, *Solid State Theory: An Introduction*, 2nd edn. (Springer, Berlin, 2009)
[10K1] C. Klingshirn et al., Phys. Status Solidi B **247**, 1424 (2010)
[10K2] C. Klingshirn, B.K. Meyer, A. Waag, A. Hoffmann, J. Geurts, *Zinc Oxide: From Fundamental Properties Towards Novel Applications*. Springer Series in Materials Science, vol. 120 (Springer, Berlin, 2010)
[10Y1] P.Y. Yu, M. Cardona, *Fundamentals of Semiconductors*, 4th edn. (Springer, Berlin, 2010)
[11S1] R. Scheer, H.-W. Schock, *Chalcogenide Photovoltaics* (Wiley-VCH, Weinheim, 2011)
[12E1] J. Even et al., Phys. Rev. B **86**, 205301 (2012)
[14G1] M.A. Green, A. Ho-Baillie, H.J. Snaith, Nat. Photonics **8**, 506 (2014)
[16A1] N.W. Ashcroft, N.D. Mermin, D. Wei, *Solid State Physics*, Revised edn. (Cengage Learning Asia, Singapore, 2016)

Chapter 16
Optical Band-to-Band Transitions

Abstract The optical constants and the dielectric function of semiconductors are over a wide range of photon energies determined by electronic transitions between various bands of the band structure. We will describe here the theoretical approach to treat these transitions arriving finally at an expression for the absorption coefficient. We will introduce the joint density of states and its critical points leading to prominent structures (van Hove-singularities) in the dielectric function. In the experimental section we will illustrate methods to determine optical functions (ellipsometry and modulation spectroscopy).

16.1 Perturbation Theory of Direct Band-to-Band Transitions

In Sect. 6.2 we have already outlined some basics of treating light-matter interaction in perturbation theory. This was discussed for the case of an ensemble of two-level atoms. The essential result was Fermi's golden rule (6.2) relating the total transition rate W_{ji} from the initial states i to the final states f of the ensemble to the matrix element of the perturbation operator $H_{fi}^{(1)}$ (transition matrix element) and the density $D(E)$ of final states at $E = \hbar(\omega_f - \omega_i)$. The perturbation operator is given in dipole approximation (see (6.23)) by the dipole operator H^D.

We will summarize here the major steps arriving at the *total transition rate for band-to-band transitions*. For a detailled discussion see e.g., [72W1, 02B1]. We start with the first-order perturbation operator $H^{(1)}$ given by (6.10):

$$H^{(1)} = -\frac{e}{m}\boldsymbol{A}\frac{\hbar}{\mathrm{i}}\nabla\,, \tag{16.1}$$

where $\boldsymbol{A}$ is the vector potential and $\boldsymbol{p} = \frac{\hbar}{\mathrm{i}}\nabla$ is the momentum operator. To calculate the transition matrix element $H_{fi}^{(1)}$ between two bands we need now the electron wavefunctions of the initial and final electron states, which we chose to be Bloch functions of the valence and conduction band, respectively:

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_16

$$\psi_i = \frac{1}{\sqrt{V}} \, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\boldsymbol{r}} \, u_{\mathrm{VB}}(\boldsymbol{r}, \boldsymbol{k}) \,, \tag{16.2a}$$

$$\psi_f = \frac{1}{\sqrt{V}} \, \mathrm{e}^{\mathrm{i}\boldsymbol{k}'\boldsymbol{r}} \, u_{\mathrm{CB}}(\boldsymbol{r}, \boldsymbol{k}') \,. \tag{16.2b}$$

The vector potential is given by a plane-wave ansatz

$$\boldsymbol{A} = \frac{1}{2} \, A_0 \, \boldsymbol{a}_0 (\mathrm{e}^{\mathrm{i}(\boldsymbol{q}\boldsymbol{r}-\omega t)} + \mathrm{e}^{-\mathrm{i}(\boldsymbol{q}\boldsymbol{r}-\omega t)}) \tag{16.3}$$

to describe absorption and stimulated emission (first and second term in the bracket, respectively). Here A_0 is the amplitude of the vector potential, a_0 the unitvector of the polarization, $\boldsymbol{q}$ and ω are the wave vector and angular frequency of the light wave.

The vector potential (16.3) is now put into (16.1) but neglecting the stimulated emission term due to lack of population in the conduction band. This perturbation operator and the Bloch functions of (16.2) are now used to determine the transition matrix element $H_{fi}^{(1)}$. We now assume a vanishing photon momentum $\boldsymbol{q} = 0$, i.e., $\boldsymbol{k} = \boldsymbol{k}'$. Due to the lattice periodicity one can substitute the integration over full space (see (6.16)) by an integration over an individual unit cell and the summation over all unit cells.

The transition matrix element reads now:

$$H_{fi}^{(1)} = -\frac{\mathrm{i}e\hbar A_0}{2m V_{cell}} \int\limits_{cell} \mathrm{d}\boldsymbol{r}' \, u_{\mathrm{LB}}^{*}(\boldsymbol{k}) \left[\boldsymbol{a}_0 \cdot \nabla u_{\mathrm{VB}}(\boldsymbol{k}) + \mathrm{i}\boldsymbol{a}_0 \cdot \boldsymbol{k} u_{\mathrm{VB}}(\boldsymbol{k}) \right] . \tag{16.4}$$

The second term in the square brackets is zero since the periodic part of the Bloch functions are orthonormal. For dipole-forbidden transitions the first term would also be zero. In this case the assumption of $\boldsymbol{q} = 0$ has to be lifted. For transitions which are allowed by symmetry we find the transition matrix element to be given by the momentum matrix element:

$$\begin{aligned} H_{fi}^{(1)} &= -\frac{\mathrm{i}e\hbar A_0}{2m V_{\mathrm{cell}}} \int\limits_{\mathrm{cell}} \mathrm{d}\boldsymbol{r}' \, u_{\mathrm{LB}}^{*}(\boldsymbol{k}) \, \boldsymbol{a}_0 \cdot \nabla u_{\mathrm{VB}}(\boldsymbol{k}) \\ &= \frac{eA_0}{2m} \, \boldsymbol{a}_0 \cdot \boldsymbol{p}_{fi} \,. \end{aligned} \tag{16.5}$$

For the transition rate between states i and f we get:

$$\begin{aligned} w_{fi}(\hbar\omega) &= \frac{2\pi}{\hbar} \left| H_{fi}^{(1)} \right|^2 \delta(E_{fi} - \hbar\omega) \\ &= \frac{\pi e^2 A_0^2}{2m^2 \hbar} \left| \boldsymbol{a}_0 \cdot \boldsymbol{p}_{fi} \right|^2 \delta(E_{fi} - \hbar\omega) \,. \end{aligned} \tag{16.6}$$

Now we want to consider the case of a simple band structure with one valence and one conduction band. The should both be parabolic and are separated by a band gap E_g. Since the absorbed photons have zero momentum each state in the valence band is coupled in the optical transition to a well defined state in the conduction band with same $\boldsymbol{k}$. The transition energy as a function of $\boldsymbol{k}$, which is called *optical energy band* is then given by $E_{CV}(\boldsymbol{k}) = E_{CB}(\boldsymbol{k}) - E_{VB}(\boldsymbol{k})$. For fixed photon energy $\hbar\omega$ we will essentially have to integrate over a surface of constant energy in reciprocal space.

Using $(4\pi^3)^{-1}$ for the density of states in $\boldsymbol{k}$-space per unit volume (in real space) including a factor of 2 for spin we find for the total transition rate for incident photons with energy $\hbar\omega$:

$$W_{CV}(\hbar\omega) = \frac{\pi e^2 A_0^2}{2m^2\hbar} \int \mathrm{d}\boldsymbol{k}\, \frac{1}{4\pi^3} \left|\boldsymbol{a}_0 \cdot \boldsymbol{p}_{CV}\right|^2 \delta(E_{CV}(\boldsymbol{k}) - \hbar\omega)\,. \tag{16.7}$$

The term $\left|\boldsymbol{a}_0 \cdot \boldsymbol{p}_{CV}\right|^2$ still depends on $\boldsymbol{k}$. But for spherical symmetry, which can be assumed approximately for cubic crystals when band warping (see Sect. 15.2.1) is neglected, this term is constant for all $\boldsymbol{k}$ forming a surface of constant energy. It then has a value averaged over all reciprocal-space dimensions $3 < p_{CV}^2 >$ and can be put outside of the integral. We also use $A_0^2 = 2\hbar/\epsilon_1\epsilon_0\omega$ to arrive at:

$$W_{CV}(\hbar\omega) = \frac{3\pi e^2 < p_{CV}^2 >}{\epsilon_1\epsilon_0\omega m^2} \int \mathrm{d}\boldsymbol{k}\, \frac{1}{4\pi^3}\, \delta(E_{CV}(\boldsymbol{k}) - \hbar\omega)\,. \tag{16.8}$$

The integral in (16.8) is called the *joint density of states* J_{CV} (actually multiplied with a factor of two to account for the spin). This quantity reflects the fact that the initial and final states for the optical transitions between the two bands are unambiguously linked by the assumption of vanishing photon momentum. For interpretation of spectroscopic data, however, it is more useful to know the joint density of states as a function of photon energy. For this we have to replace the integration over $\boldsymbol{k}$ by an integration over a surface of constant energy and find with (4.15):

$$D(E_{CV}) = \frac{1}{4\pi^3} \int \frac{\mathrm{d}S}{|\nabla_{\boldsymbol{k}} E_{CV}|_{E_{CV}=\hbar\omega}} = 2J_{CV}(\hbar\omega)\,. \tag{16.9}$$

To calculate the total transition rate we now use the fact, that the individual density of states of the parabolic valence and conduction bands each have a square-root dependence on energy. From

$$E_{CV} = E_g + \frac{\hbar^2 k^2}{2\mu} \qquad \text{with} \quad \frac{1}{\mu} = \frac{1}{m_e} + \frac{1}{m_h} \tag{16.10}$$

we finally get:

$$W_{CV}(\hbar\omega) = \frac{3e^2(2\mu)^{3/2} < p_{CV}^2 >}{2\pi\hbar^3 m^2 \epsilon_1\epsilon_0\omega} (\hbar\omega - E_g)^{1/2}\,. \tag{16.11}$$

Now using the definition of the average oscillator strength in momentum space

$$< f >= 2 < p_{\mathrm{CV}}^2 > /m\,\hbar\omega \tag{16.12}$$

we can write:

$$W_{\mathrm{CV}}(\hbar\omega) = \frac{3(2\mu)^{3/2}}{4\pi\hbar^2\epsilon_1\omega}\frac{e^2}{m\epsilon_0} < f > (\hbar\omega - E_{\mathrm{g}})^{1/2}\,. \tag{16.13}$$

We see that the total transition rate contains the reduced mass, the classical oscillator strength (second fraction in (16.13)), and the average oscillator strength of the band-to-band transition and it depends (like the joint density of states) on the square-root of the photon energy exceeding the gap energy.

The energy dependence of the joint density of states can be deduced from comparison of (16.8) and (16.13) to be:

$$J_{\mathrm{CV}}(\hbar\omega) = \frac{3(2\mu)^{3/2}}{2\pi^2\hbar^3}(\hbar\omega - E_{\mathrm{g}})^{1/2}\,. \tag{16.14}$$

This joint density of states $J_{\mathrm{CV}}(\hbar\omega)$ is now given in units of $\mathrm{eV}^{-1}\,\mathrm{cm}^{-3}$.

As final result of this discussion of band-to-band transition we will now deduce a measurable quantity, namely the *absorption coefficient*. This coefficient is defined in this microscopic treatment by the ratio of energy dissipation (transition rate times absorbed energy per transition) and energy flux density (energy density times energy velocity):

$$\alpha_{\mathrm{fi}}(\hbar\omega = \hbar\omega_{\mathrm{fi}}) = \frac{\hbar\omega_{\mathrm{fi}} W_{\mathrm{fi}}}{\epsilon_1\epsilon_0 < \boldsymbol{E}^2 > v_{\mathrm{g}}} \tag{16.15}$$

Using energy quantization $\epsilon_1\epsilon_0 < \boldsymbol{E}^2 > = \hbar\omega$ and the relation $\epsilon_1 v_{\mathrm{g}} = nc$ we find with (16.13) for band-to-band absorption:

$$\alpha_{\mathrm{CV}}(\hbar\omega) = \frac{\epsilon_1 W_{\mathrm{CV}}}{nc} = \frac{3e^2(2\mu)^{3/2} < f >}{4\pi\hbar^2 nc\epsilon_0 m}(\hbar\omega - E_{\mathrm{g}})^{1/2}\,. \tag{16.16}$$

Since the energy dependence of $< f >$ is negligible for energies close to the band gap we see that the absorption is a square-root function starting at E_{g} reflecting the joint density of states.

The absorption coefficient for direct band-to-band transitions is displayed in Fig. 16.1 together with the absorption at an indirect band gap. We notice that the absorption across an indirect band gap leads to quadratic dependencies on photon energy.

The absorption across an indirect gap like in Si or Ge involves the absorption and emission of phonons to fulfill momentum conservation. The theoretical treatment of these transitions requires perturbation theory of second order (see [02B1]). The involved phonons have to fulfill group theoretical selection rules. The absorption-

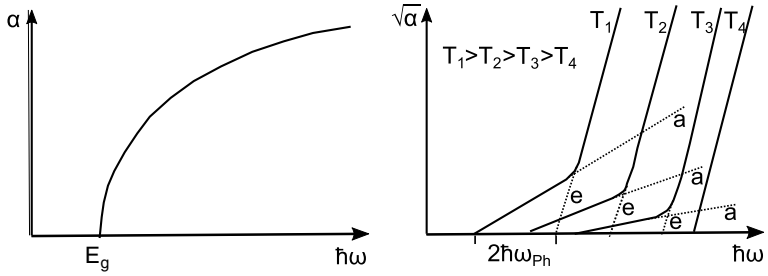


Fig. 16.1 Absorption coefficient at a direct band gap (**a**) and an indirect band gap (**b**). In the later absorption (**a**) and emission (**e**) of phonons contribute to the optical absorption process

assisted processes start one phonon energy below the energy of the band gap while the ones involving phonon emission start one phonon energy above. Due to the bosonic character of the phonons the temperature-dependent number density of phonons in the semiconductor enters into the transition rate of the second-order processes both for emission and absorption of phonons. This is particularly obvious for the vanishing contribution of phonon absorption for low temperatures (see Fig. 16.1b).

A final comment is necessary here. The realistic absorption spectra of semiconductors at the fundamental gap are much more complex. For indirect-gap materials different types of phonons can contribute (like LO and LA phonons in Ge, or GaAs-type and AlAs-type phonons in indirect-gap $Al_{1-y}Ga_yAs$). Further, effects due to the Coulomb interaction between electron and hole (excitons and enhanced oscillator strength) strongly modify the absorption spectra. We will discuss these effects in detail later in Chap. 21.

16.2 Optical Functions of Semiconductors and Their Critical Points

So far we have concentrated our discussion on the fundamental band gap. We now want to consider the optical properties of semiconductors over a large energy range from zero to photon energies beyond the highest electronic transitions. In particular we want to focus on the real and imaginary part of the dielectric function, which can be measured e.g., by ellipsometry (see Fig. Ell.1) . We recall the relations

$$\alpha_{\mathrm{fi}}(\omega) = \frac{\epsilon_1(\omega) W_{\mathrm{fi}}(\omega)}{n(\omega) c} = \frac{2\omega}{c}\kappa(\omega) \qquad \text{and} \tag{16.17a}$$

$$\epsilon_2(\omega) = 2n(\omega)\kappa(\omega) \tag{16.17b}$$

and find with (6.17):

$$\epsilon_2(\omega) = \frac{e^2}{4\pi^2 \epsilon_0 m_0^2 \omega^2} \int \mathrm{d}\boldsymbol{k} \; \left| \boldsymbol{a}_0 \cdot \boldsymbol{p}_{\mathrm{fi}} \right|^2 \, \delta(E_{\mathrm{fi}}(\boldsymbol{k}) - \hbar\omega) \, . \tag{16.18}$$

We use the definition of the joint density of states of (16.9) and keep in mind that the term $\left| \boldsymbol{a}_0 \cdot \boldsymbol{p}_{\mathrm{fi}} \right|^2$ varies only slightly in frequency areas where $J_{\mathrm{fi}}(\omega)$ shows strong changes (e.g., at the band gap). With this we arrive at

$$\epsilon_2(\omega) = \frac{2\pi e^2}{\epsilon_0 m_0^2 \omega^2} \left| \boldsymbol{a}_0 \cdot \boldsymbol{p}_{\mathrm{fi}} \right|^2 \cdot J_{\mathrm{fi}}(\omega) \, . \tag{16.19}$$

Again we find a direct relation between an optical function and the joint density of states of the related optical energy band.

Inspecting (16.9) we expect pronounced structures in $J_{\mathrm{fi}}(\omega)$ and thus $\epsilon_2(\omega)$ for points in $\boldsymbol{k}$-space where the derivative of the optical energy band $\nabla_{\boldsymbol{k}} E_{\mathrm{fi}}$ vanishes. These critical points are called *van Hove singularities*. They occur for an optical energy band if

$$\nabla_{\boldsymbol{k}} E_{\mathrm{fi}} = \nabla_{\boldsymbol{k}} E_{\mathrm{f}} - \nabla_{\boldsymbol{k}} E_{\mathrm{i}} = 0 \, , \tag{16.20}$$

i.e., in case both derivatives are zero or otherwise exactly equal. The first case occurs often at points of high symmetry in the electronic band structure (e.g., at the band gap of a direct semiconductor). The later case occurs when the two electronic bands (typically by accident) have the same slope in $\boldsymbol{k}$-space. The related abrupt changes in the energy dependence of J_{fi} are called minima (M_0) and maxima (M_3), or saddle points (M_1) and (M_2), respectively (see Fig. 16.2).

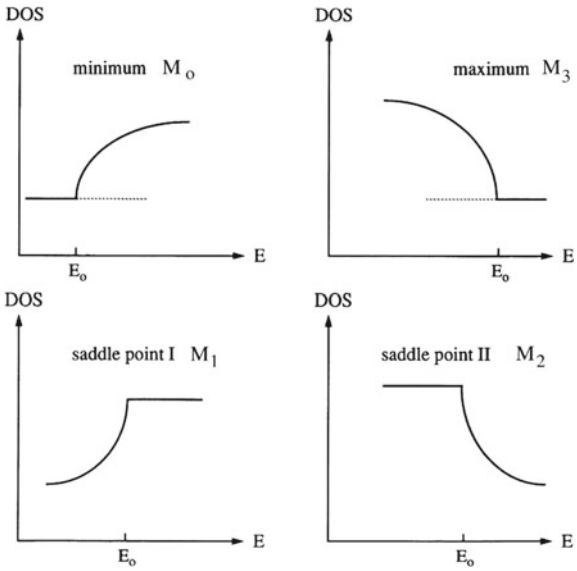


Fig. 16.2 Schematic drawing of critical points in the joint density of states of an optical energy band

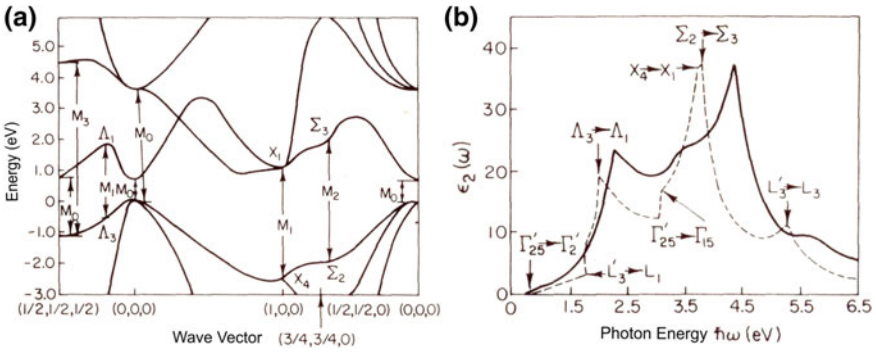


Fig. 16.3 Calculated band structure of Ge indicating the position of critical points in the joint density of states (**a**). Calculated (solid line) and measured (dashed line) imaginary part of the dielectric function $\epsilon_2(\omega)$ (**b**) (adapted using the notation of [72W1] with permission from [62B1] ©1962 American Physical Society https://doi.org/10.1103/PhysRevLett.9.94

Experimental determination of the critical points in $\epsilon(\omega)$ is very useful for the detailed elaboration of the band structure of the respective semiconductor. It provides useful input parameters for band-structure calculations. This is demonstrated in Fig. 16.3 where we show a calculated Ge band structure in **(a)** in comparison to respective calculated and the measured imaginary part of the dielectric function $\epsilon_2(\omega)$ in **(b)**.

The pronounced features in $\epsilon_2(\omega)$ can directly by correlated with critical points in the joint density of states as indicated in the band structure. The indirect bad gap does not affect much the dielectric function. Its onset is dominated by the direct gap at the Γ-point. Much stronger features are due to saddle points. Please note that the peak at about 4.5 eV actually results from the superposition of two saddle points.

This brings us to the point where we need to discuss how to measure the optical functions, in particular over a broad spectral range of several eV. It easy to understand that the refractive index of a sample can readily be determined for photon energies below the band gap with the use of a refractometer or an interferometer. Above the band-gap energy one can determine the absorption coefficient by transmission experiments, but only for very thin slaps of semiconductors.

Better accessible is the reflectivity $R(\omega)$ of semiconductor samples which, however, requires flat and non-contaminated surfaces. The spectral dependence can be measured using broad band discharge lamps or a synchrotron. The later provides radiation of high brilliance from the infrared to the X-ray regime. One can derive Kramers–Kronig relations (see Appendix A) for the amplitude and phase of the complex reflection for normal incidence (A.13). From a measurement the reflectivity for a large frequency range it is thus possible to determine the spectral dependence of the dielectric function or the refractive index (for more details see [10Y1]).

When detecting the reflectivity for fixed-angle oblique incidence of light polarized parallel and perpendicular to the plane of incidence, one can deduce the complex refractive index from Fresnel's equations (5.18) [10Y1]. Widely used is a

modification of this oblique-angle reflectivity experiment, which is called *ellipsometry* (see Fig. Ell.1).

The reflection spectra (see Fig. Ell.2) often show rather smooth contours close to minima or maxima of the joint density of states. In order to determine optical transitions related to the fundamental or higher-lying band gaps, i.e., the energies of the gaps or of excitonic resonances, with high precision one can use modulation-spectroscopy techniques Fig. Ell.2.

Experimental Technique: Ellipsometry

The measurement of the polarization change of elliptically polarized light due to reflection from a solid's surface is called *ellipsometry*. This technique allows to determine simultaneously the real and imaginary part of the dielectric function. In layered structures also the thickness of a thin film can be precisely determined. For reviews see, e.g., [04S1, 09K1]

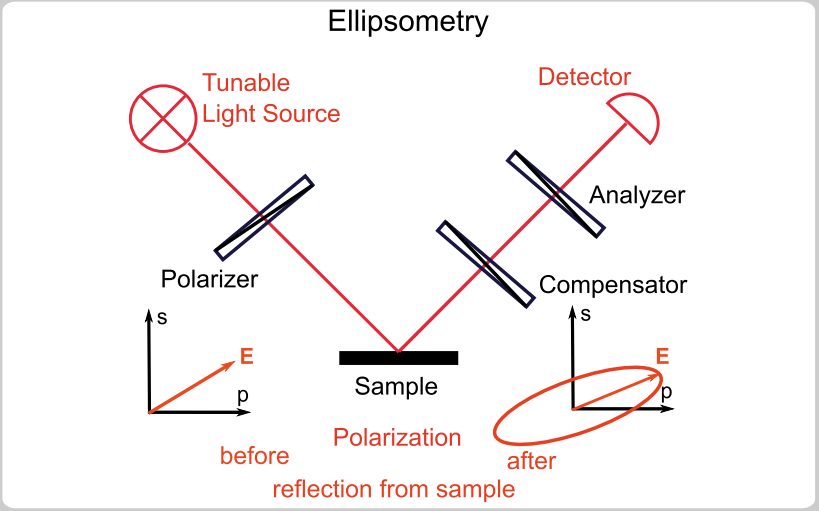


Fig. Ell.1 Schematic set-up of an ellipsometry experiment. The tunable light source is either a narrow-band emitter with a wide tunability or a broad-band emitter followed by a monochromator. The polarization state of the light before and after hitting the sample is indicated

Light from a tunable light source is linearly (neither s- nor p-) polarized by a polarizer before impinging under oblique angle ϕ on the sample surface. This light is polarized elliptically after reflection from the sample due to a phase shift between the s- and p-polarization components. This phase shift can be determined using the compensator which introduces an additional variable phase shift to extinct the signal in combination with the analyzer. Alternatively,

the analyzer is rotated with fixed angular frequency leading to an oscillating signal at the detector. In both cases one determines the orientation and axial ratio of the ellipse which yields the ratio of the complex reflection coefficients $\sigma = r_p/r_s$. The complex dielectric function can then be deduced from [10Y1]:

$$\epsilon = \sin^2\phi + \sin^2\phi \tan^2\phi \left(\frac{1-\sigma}{1+\sigma}\right)^2 . \qquad \text{(Ell-1)}$$

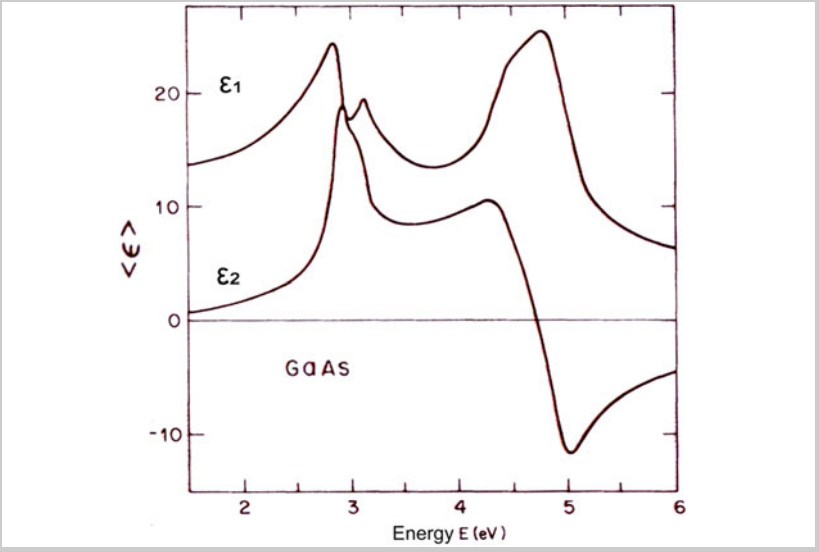


Fig. Ell.2 Dielectric function of GaAs determined from ellipsometry (reprinted with permission from [83A1] ©1983 American Physical Society https://doi.org/10.1103/PhysRevB.27.985)

Experimental Technique: Modulation Spectroscopy

Inspecting the spectra of the joint density of states in Fig. 16.2 reveals that the derivative of $J_{fi}(\omega)$ diverges at critical point (or should have at least distinct behavior in realistic, i.e., broadened spectra). The basic idea of *modulation spectroscopy* techniques [94P1, 89C1, 10Y1] is to exploit the high sensitivity to critical points when measuring the signal derivative. This can be done by periodic modulation of the electronic states coupled with phase-synchronous detection. Measuring the respective signal change (e.g., change in reflectivity ΔR) normalized to the signal ($\Delta R/R$) leads to a self-referenced measurement independent of spectral changes of the detector sensitivity or the incident light

intensity. The modulation can be induced by an AC electric field (*electro-reflectance (ER)*, modulated illumination (*photo-reflectance*), modulated strain (*piezo-reflectance*) etc. Of course, the experiments can in principle also be done in transmission geometry (Fig. Mod.1).

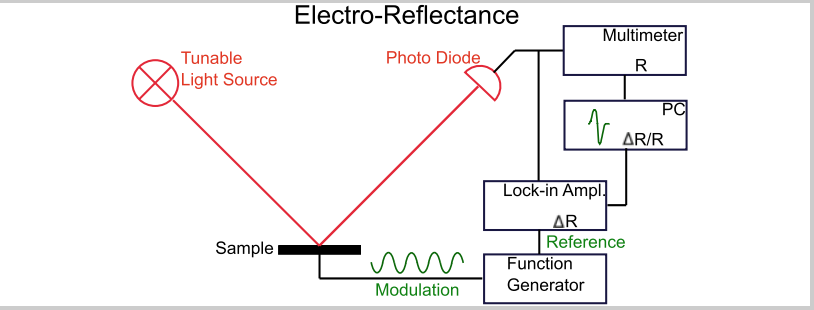


Fig. Mod.1 Schematic set-up of an electro-reflectance experiment. Here the modulation is given by an AC voltage applied to the sample. The electro-reflectance spectrum $\Delta R/R$ is calculated from the integrated reflectivity R and phase-synchronous derivative ΔR

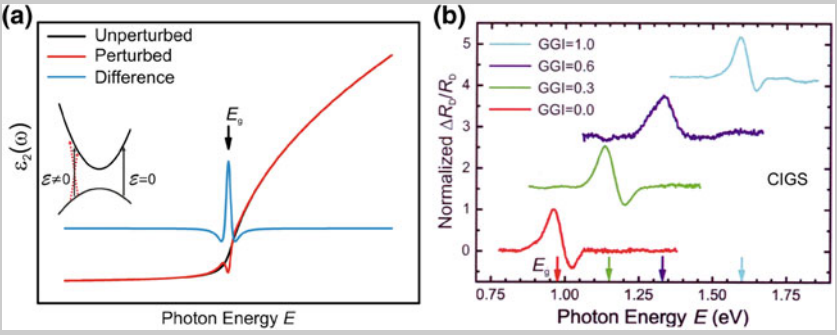


Fig. Mod.2 Imaginary part of the dielectric function $\epsilon_2(\omega)$ of a semiconductor with (red) and without (black) applied electric field [15K1]. The difference $\Delta\epsilon_2$ is centered around the fundamental bandgap and vanishes for photon energies away from the resonance**(a)** and diffuse ER spectra of CIGS solar-cells absorber layers for increasing Ga content **(b)**. The band-gap energies E_g indicated by the arrows in (b) have been determined from the modulus spectra [15K1]

Modulation spectroscopy has a manifold of applications. It was heavily used to determine the band-structure details of the common semiconductors in the past (see, e.g., [73A1, 88P1, 89C1, 94P1]), and is of course still attractive for novel materials or simply to characterize samples after fabrication.

Panel (a) of Fig. Mod.2 show the modulation of the dielectric function upon application of an electric field. The resonance-like feature close to the fundamental gap reveal the transition energy between valence and conduction band. This also applies to transitions between lower valence and/or higher

conduction bands (see, e.g., [73A1]). The modulation of the dielectric function is directly reflected in the ER spectra. These spectra actually need a careful line-shape analysis involving fits of derivative functions. For the situation of freely moving carriers in the sample a third derivative functional form (TDFF) is applied [73A2], while a first derivative needs to be considered in case of bound transitions [88P1]. The later is the case for pronounced excitonic resonances which actually dominate the spectra of many semiconductors [88P1]. Further, Franz–Keldysh oscillations on the high- energy side of the spectra become increasingly important for high applied electric fields [92G1].

The experimental spectra in Fig. Mod.2 are actually taken on samples which are not high-quality, single-crystal semiconductors with perfect surfaces. This is the case, e.g., for thin-film solar cells, where the band gap of a poly-crystalline, alloy absorber layer needs to be determined. The figure shows shows ER spectra taken on $CuIn_{1-x}Ga_xSe_2$ solar cells with increasing Ga content in the absorber layer (GGI is the Ga to Ga+In ratio). Strong interference effects due to the layered structure, which dominate the specular reflection, have been suppressed by recording diffuse ER [15K1, 15K2]. In principle, the band-gap energy can be deduced from a TDFF fit. The evaluation of the gap energy is, however, easier from the modulus spectrum. Since the measured signal can be considered as real part of a complex optical function, one can determine the imaginary part by Kramers–Kronig transformation. The calculated modulus of the complex spectrum has a peak form where the maximum directly indicates the position of the critical point [94H1].

16.3 Occupation of Electronic Band States

So far the occupation of electronic states in the bands did not play a significant role, except for the assumption that in the case of band-to-band absorption (see (16.16) and Fig. 16.1) the initial states are occupied while the final states are empty. But for finite temperatures the population of the bands is non-zero and can achieve large values in the case of doped semiconductors. Further, optical excitation or electrical injection drive the semiconductor into a non-equilibrium situation by generating excess carriers in the bands.

Since electrons are fermions the occupation probability of carriers in the bands is given by *Fermi–Dirac statistics*. It reads for electrons and holes

$$\text{electrons}: \; f_{\mathrm{FD}}(E) = \left(\mathrm{e}^{(E-E_{\mathrm{F}}^{\mathrm{e}})/(k_{\mathrm{B}}T)} + 1\right)^{-1} \tag{16.21a}$$

$$\begin{aligned}\text{holes}: \; f_{\mathrm{FD}}(E) &= 1 - \left(\mathrm{e}^{(E-E_{\mathrm{F}}^{\mathrm{h}})/(k_{\mathrm{B}}T)} + 1\right)^{-1} \\ &= \left(\mathrm{e}^{-(E-E_{\mathrm{F}}^{\mathrm{h}})/(k_{\mathrm{B}}T)} + 1\right)^{-1}.\end{aligned} \tag{16.21b}$$

The respective Fermi energies (or chemical potentials) for electrons and holes $E_{\mathrm{F}}^{\mathrm{e,h}}$ are non-zero, unlike the situation for phonons or photons, since there is a conservation law for the number of electrons. The Fermi energies depend on the concentrations of electrons and holes, on the temperature of the carrier gas, and on material parameters such as E_{g}, m_{e} or m_{h}.

Since electrons can be exchanged between the valence and conduction bands, for example by thermal excitation and by recombination, it follows that in thermodynamic equilibrium

$$E_{\mathrm{F}}^{\mathrm{e}} = E_{\mathrm{F}}^{\mathrm{h}} = E_{\mathrm{F}} \tag{16.22a}$$

or, in other words, the chemical potential of the electron–hole pair system μ_{eh} is zero

$$\mu_{\mathrm{eh}} = E_{\mathrm{F}}^{\mathrm{e}} - E_{\mathrm{F}}^{\mathrm{h}} = 0\,. \tag{16.22b}$$

The density of electrons in a certain energy range, say in the conduction band, is then given by the integral over the density of states weighted by the occupation probability. For parabolic bands with a degeneracy g_{e} we find for the total density of electrons in the conduction band n:

$$n = \frac{N}{L^3} = \int_{E_{\mathrm{g}}}^{\infty} g_{\mathrm{e}}\,\frac{1}{2\pi^2}\left(\frac{2m_{\mathrm{e}}}{\hbar^2}\right)^{3/2}\left(E - E_{\mathrm{g}}\right)^{1/2}\left(\mathrm{e}^{(E-E_{\mathrm{F}})/(k_{\mathrm{B}}T)} + 1\right)^{-1}\mathrm{d}E\,, \tag{16.23}$$

where we have used our knowledge from Chap. 4 and simple parabolic dispersion relations in effective-mass approximation for the bands:

$$\text{conduction band (CB)}: \quad E(\boldsymbol{k}_{\mathrm{e}}) = E_{\mathrm{g}} + \frac{\hbar^2\boldsymbol{k}_{\mathrm{e}}^2}{2m_{\mathrm{e}}}, \tag{16.24a}$$

$$\text{valence band (VB)}: \quad E(\boldsymbol{k}_{\mathrm{h}}) = \frac{\hbar^2\boldsymbol{k}_{\mathrm{h}}^2}{2m_{\mathrm{h}}}\,. \tag{16.24b}$$

Equation (16.23) leads to the *Fermi integral* which can only be solved numerically.

Very often E_{F} lies in the band gap, i.e., $E_{\mathrm{F}} < E_{\mathrm{g}}$. In this case the electron gas is said to be *non-degenerate* and the part of the Fermi function which overlaps with a finite density of states can be approximated quite well by the Boltzmann function. This has been illustrated in Fig. 4.3 where we compared Fermi–Dirac, Bose–Einstein and Boltzmann statistics. With this assumption the Fermi–Dirac statistics (16.23) can be replaced by classical Boltzmann statistics and the integral can be solved analytically giving

$$n = g_{\mathrm{e}}\left(\frac{m_{\mathrm{e}}k_{\mathrm{B}}T}{2\pi\hbar^2}\right)^{3/2}\mathrm{e}^{-(E_{\mathrm{g}}-E_{\mathrm{F}})/(k_{\mathrm{B}}T)}. \tag{16.25a}$$

The term in front of the exponential is the so-called *effective density of states* n_{eff}, which depends on T and on material parameters and describes the capacity of the

band to accommodate carriers. So we can write

$$n = n_{\text{eff}}(T, m_{\text{e}})\, \mathrm{e}^{-(E_{\text{g}}-E_{\text{F}})/(k_{\text{B}}T)}. \tag{16.25b}$$

If the electron gas is *degenerate*, i.e., if E_{F} is situated within the conduction band, we have to use the full (16.23) except for $T \to 0$ where the Fermi function converges to a step function and n is given by

$$n = \frac{2}{6\pi^2}\, g_{\text{e}} \left(\frac{2m_{\text{e}}}{\hbar^2}\right)^{3/2} \left(E_{\text{F}} - E_{\text{g}}\right)^{3/2} \quad \text{for} \quad T = 0\,\text{K}\,. \tag{16.25c}$$

In analogy we find for a three-dimensional non-degenerate gas of holes

$$p = p_{\text{eff}}\, \mathrm{e}^{-E_{\text{F}}/k_{\text{B}}T} \tag{16.25d}$$

with

$$p_{\text{eff}} = g_{\text{h}} \left(\frac{m_{\text{h}} k_{\text{B}} T}{2\pi\hbar^2}\right)^{3/2}. \tag{16.25e}$$

In thermodynamic equilibrium, i.e., when (16.22a) is valid, one finds that the carrier-density product np is independent of E_{F} in the non-degenerate case:

$$np = n_{\text{eff}}\, p_{\text{eff}}\, \mathrm{e}^{-E_{\text{g}}/k_{\text{B}}T} =: n_{\text{i}}^2(T)\,. \tag{16.26}$$

where $n_{\text{i}}(T)$ is the *intrinsic carrier density*. Equation (16.26) is called the *mass-action law* of semiconductor physics. It also applies to doped materials. This means that increasing the density of one carrier type by some orders of magnitude by doping results in a respective reduction of the other type.

To get a feeling for relevant orders of magnitude we consider the example of GaAs at $T = 300$ K. The conduction band is spin-degenerate and has an effective mass of $m_{\text{e}} = 0.067 m_0$. For the valence band one has to consider the population of heavy- and light-hole bands (both spin-degenerate) which is taken care of by the averaged effective hole mass $m_{\text{h}} = (m_{\text{lh}}^{3/2} + m_{\text{hh}}^{3/2})^{2/3}$. For GaAs we find $m_{\text{h}} = 0,47 m_0$. The effective densities of states are then $n_{\text{eff}} = 4.2 \cdot 10^{17}\,\text{cm}^{-3}$ and $p_{\text{eff}} = 9.4 \cdot 10^{18}\,\text{cm}^{-3}$. The intrinsic carrier density at room temperature is comparably tiny with $n_{\text{i}} \approx 5 \cdot 10^7\,\text{cm}^{-3}$.

The occupation of donor or acceptor levels situated in the forbidden gap or of localized tail states (both scenarios will be covered in detail in Chap. 18) can also be covered by the concept of Fermi–Dirac occupation probabilities. For defect states degeneracy factors appear in the Fermi statistics which take care, e.g., of the fact that an unoccupied simple donor in a direct gap semiconductor offers two empty, spin-degenerate states. But if one electron occupies the donor state, the other electron level is shifted to higher energies due to Coulomb interaction. For details of this topic see the discussion of (18.4) or [87B1].

Out of thermal equilibrium, a situation that occurs, e.g., under optical pumping or via carrier injection in a *p-n* junction, one can assume frequently a thermalized distribution of the carriers within their respective bands. The state population can be described by some effective carrier temperature $T_{e,h}$, which may be higher than the lattice temperature T_L. In this case the carrier distribution can be still described by Fermi-functions, which contain separate *quasi-Fermi energies* for electrons E_F^e and holes E_F^h. In such non-equilibrium situations the mass action law does not apply, i.e.,

$$n \cdot p \neq n_i^2 \ (T) \tag{16.27a}$$

and in contrast to (16.22b) the chemical potential of the electron–hole pair system is now finite

$$\mu_{eh} = E_F^e - E_F^h \neq 0\,. \tag{16.27b}$$

The carrier system of a semiconductor which has been brought into a *non-equilibrium distribution* will always tend to relax towards equilibrium. If the carrier product $n \cdot p$ deviates from the product in equilibrium given by the mass action law (16.26) the system will react with net recombination (or generation, respectively) of carriers. Such recombination processes can be radiative and are then detectable as luminescence. This situation is exploited in spectroscopic characterization of semiconductors (photoluminescence, electroluminescence) and in devices like light-emitting or laser diodes. If carriers are generated with excess energy with respect to the band minimum (for the case of electrons) or maximum (holes) the carrier population will rapidly thermalize to establish a carrier temperature. This is followed by a cooling of the carrier distribution towards the sample temperature. These relaxation processes are dominated by interaction with the phonon system of the semiconductor.

In the following, we will frequently use data from luminescence spectroscopy for characterization of the electronic states. Luminescence is a form of linear spectroscopy since optical functions of the semiconductor are not affected by the non-equilibrium carrier population (at least for low excitation levels). All other dynamical processes evoked due to a non-equilibrium situation will be treated in detail in the second volume of this text book.

Problems

16.1 Calculate the effective density of states, i.e., the onset of degeneracy, for electrons and holes at 10 K and at room temperature in bulk GaAs and ZnSe.

16.2 Calculate the effective density of states for the case of electrons and holes in Si and Ge at room temperature. Pay attention to the presence of equivalent conduction-band minima and find a suitable scheme to average over transverse and longitudinal masses.

16.3 Verify the type of the different critical points in Fig. 16.3a and try to construct qualitatively the imaginary part of the dielectric function for Ge as shown in (b). Try the same for GaAs by comparison of Fig. Ell.2 with Fig. 15.8.

References

[62B1] D. Brust, J.C. Phillips, F. Bassani, Phys. Rev. Lett. **9**, 94 (1962)
[72W1] F. Wooten, *Optical Properties of Solids* (Academic, New York, 1972)
[73A1] D.E. Aspnes, A.A. Studna, Phys. Rev. B **7**, 4605 (1973)
[73A2] D.E. Aspnes, Surf. Sci. **37**, 418 (1973)
[83A1] D.E. Aspnes, A.A. Studna, Phys. Rev. B **27**, 985 (1983)
[87B1] J.S. Blakemore, *Semiconductor Statistics* (Dover Publication, Dover, 1987)
[88P1] F.H. Pollak, O.J. Glembocki, Proc. SPIE **0946**, 1 (1988)
[89C1] M.L. Cohen, J.R. Chelikowsky, *Electronic Structure and Optical Properties of Semiconductors*, vol. 75, 2nd edn., Springer Series in Solid State Sciences (Springer, New York, 1989)
[92G1] J.M.A. Gilman, A. Hamnett, R.A. Batchelor, Phys. Rev. B **46**, 13363 (1992)
[94H1] T.J.C. Hosea, Phys. Status Solidi B **182**, 43 (1994)
[94P1] F. Pollak, Modulation spectroscopy of semiconductors and semiconductor microstructures. Optical Properties of Semiconductors, vol. 2. *Handbook on Semiconductors* (North Holland, Amsterdam, 1994)
[02B1] P.K. Basu, *Theory of Optical Processes in Semiconductors: Bulk and Microstructures* (Oxford University Press, Oxford, 2002)
[04S1] M. Schubert, *Infrared Ellipsometry on Semiconductor Layer Structures* (Springer, Berlin, 2004)
[09K1] H. Kuzmany, *Solid State Spectroscopy*, 2nd edn. (Springer, Berlin, 2009)
[10Y1] P.Y. Yu, M. Cardona, *Fundamentals of Semiconductors*, 4th edn. (Springer, Heidelberg, 2010)
[15K1] C. Krämmer, *Optoelectronic Characterization of Thin-Film Solar Cells by Electroreflectance and Photoluminescence Spectroscopy*. Dissertation, Karlsruhe Institute of Technology KIT (2015)
[15K2] C. Krämmer, C. Huber, A. Redinger, D. Sperber, G. Rey, S. Siebentritt, H. Kalt, M. Hetterich, Appl. Phys. Lett. **107**, 222104 (2015)

Chapter 17
Low-Dimensional Semiconductor Structures

Abstract Low-dimensional semiconductor structures are nowadays the centerpiece of many electronic and optoelectronic devices. This is a result of their interesting properties like carrier confinement and localization, tunability of electronic energies by geometric dimension, or a spectrally narrow and high density of states. But low-dimensional structures are also perfect model systems for quantum mechanics and lead to new physics phenomena like the quantum Hall effect. We will discuss here the electronic states and wavefunctions in (square-well) potentials of different dimensionalities. We will address the consequences of finite size and depth of the potentials and how the bandstructure characteristics affect the electronic states. Starting with single quantum wells (QWs) we proceed via coupled QWs to superlattices where miniband formation leads to transport perpendicular to the quantization layers. We focus in particular on a novel class of materials—mono-layer semiconductors—like graphen and transition-metal dichalcogenides. These materials show extraordinary electronic properties due to location of their bandgaps at the corners of a hexagonal Brillouin zone leading to the phenomena of valleytronics. We then describe electronic states in quantum wires, rods and dots and how to realize such systems.

17.1 Density of States and Critical Points in Various Dimensions

A *low-dimensional electron system* can be realized, if the motion of the carriers is confined in one or more dimensions by a suitable potential on a length scale comparable to or smaller than the de Broglie wavelength of these particles, or the mean distance between scattering events in the sense of a relaxation-time approach. For semiconductors this limit is reached for the width of the confining potential typically below a few tens of nm at least at low temperatures. Since this implies a finite spatial extend of the confining potential such structures are called *quasi*-two dimensional etc. In the remaining dimensions normal to this confining potential the carriers are assumed to move freely as Bloch waves or as plane waves in the sense of the effective-mass approximation.

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_17

Start from simple parabolic bands for electrons and holes (see (16.24)) and with the help of (4.19) we find the density of states $D(E)$ of conduction and valence bands

$$\text{CB}: D(E) = \left(E - E_\text{g}\right)^{\frac{d}{2}-1} \qquad \text{for } E > E_\text{g} \tag{17.1a}$$

$$\text{VB}: D(E) = E^{\frac{d}{2}-1} \qquad \text{for } E > 0\,, \tag{17.1b}$$

for three-, two- and one-dimensional systems where $d =$ dimensionality. Please note that for holes in valence bands we assumed that the energy rises for the hole situated deeper in the band. The different density of states are shown schematically in Fig. 17.1 where we also include a set of δ-functions for a quasi-zero dimensional system.

The appearance of the various subbands for $d < 3$ will be explained in the following sections.

Also the critical points of a band structure have a characteristic dependence on dimensionality of the carrier system. We illustrated the importance of critical points in the optical properties of vertical band-to-band transitions in Sect. 16.2. In this case not the density of states in each of the two bands is important, but rather the joint density of states of conduction and valence band separated in their $E(\boldsymbol{k})$ dispersion vertically by the photon energy $\hbar\omega$.

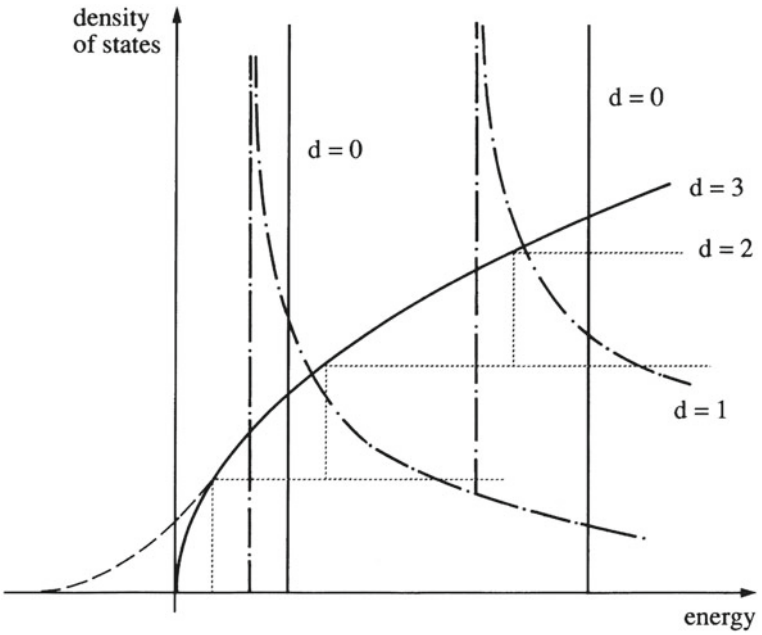


Fig. 17.1 Schematic drawing of the density of states as a function of energy for three-, and quasi-two-, one- and zero-dimensional systems in the effective-mass approximation. The dashed line corresponds to disorder-induced localized states (Sect. 18.2)

One type of these critical points occurs if a parabolic band starts at energy E_0 (e.g., a conduction band). The density of states then given by (17.1a) and zero below, i.e., in a d dimensional system

$$D(E) \sim \begin{cases} (E - E_0)^{\frac{d}{2}-1} & \text{for } E \geq E_0 \\ 0 & \text{for } E < E_0 \end{cases}. \tag{17.2a}$$

This is a so-called M_0 critical point in three dimensions.

If a parabolic band ends at a certain energy E_0 like a valence band, the inverse situation occurs, i.e.,

$$D(E) \sim \begin{cases} (E_0 - E)^{\frac{d}{2}-1} & \text{for } E \leq E_0 \\ 0 & \text{for } E > E_0 \end{cases}. \tag{17.2b}$$

The name of this type of critical point is M_{d} where d is the dimensionality.

In three dimensions there is additionally the possibility that the curvature of $E(\boldsymbol{k})$ is positive in two orthogonal directions and negative in the third one or vice versa. These so-called M_1 and M_2 critical points exhibit a constant density of states D_0 (similar to a two-dimensional system) on the energy side of E_0 where two branches exist and a dependence

$$D(E) \sim \left(1 - A\,(E - E_0)^{1/2}\right) \tag{17.2c}$$

on the other side.

A saddle point or M_1 critical point in a two dimensional system finally has a logarithmic singularity on both sides of E_0

$$D(E) \sim -\ln\left(|E - E_0|\right). \tag{17.2d}$$

For illustrations of the energy-dependence of saddle points compare to Fig. 16.2. For details of the mathematical description see, e.g, [72W1, 10Y1].

17.2 Quantum Wells and Superlattices

One of the first realizations of a quasi two-dimensional system was via simple heterostructures, where two different semiconductors where epitaxially grown on top of each other (for details on the epitaxial growth of low-dimensional semiconductors see Appendix D). If the band-discontinuity at the heterointerface and the position of the Fermi levels are suitable, an arrangement of the bottom of the conduction band may appear as depicted schematically in Fig. 17.2.

In thermodynamic equilibrium the Fermi energy is constant across the interface. The bottom of the conduction band forms a roughly triangular potential (of finite

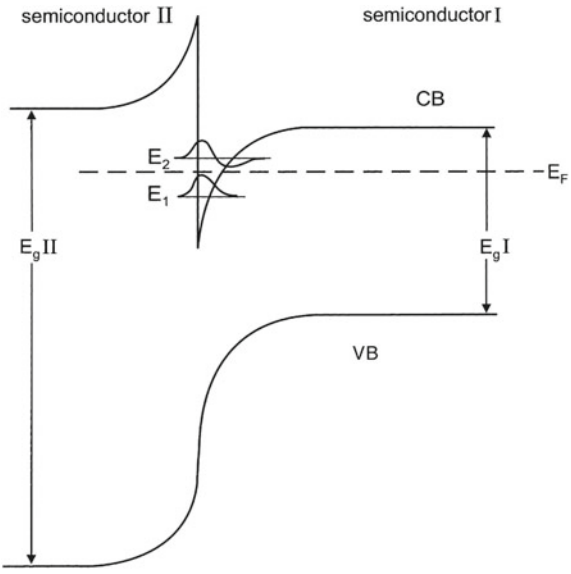


Fig. 17.2 Conduction and valence band edges around a heterojunction between two different semiconductors, forming a quasi-two dimensional electron gas at the interface. E_F: Fermi level

height only) for the electrons. The electrons can move freely in the plane of the interface but are quantized in their motion normal to it. The envelope functions of the first two quantized states are shown schematically.

In effective-mass approximation the wave function can be written as

$$\Psi_n\left(\boldsymbol{r}, \boldsymbol{k}_\parallel\right) = \mathrm{e}^{\mathrm{i}(k_x x + k_y y)} \chi_n(z) \tag{17.3a}$$

with

$$E_n\left(\boldsymbol{k}_\parallel\right) = E_n + \frac{\hbar^2 (k_x^2 + k_y^2)}{2m_\mathrm{e}} \,. \tag{17.3b}$$

The so-called envelope function $\chi^n(z)$ with quantization-energy E_n is obtained as the solution of the given potential. For the solution of the potential of Fig. 17.2 the first two quantized states are shown schematically. At every value E_n starts a quasi-two dimensional band with constant density of states as already discussed in Sect. 17.1 and shown in Fig. 17.1.

In the situation of Fig. 17.2 it is assumed that E_F is situated slightly above E_1 so that a degenerate, quasi-two-dimensional electron gas exists in the first subband already without excitation or carrier injection. Note that this structure does not provide any confining potential for the holes.

Single heterostructures like the one of Fig. 17.2 play an important role for electronic devices like field-effect transistors. But in order to exploit quantization effects, especially in (electro-)optic devices, more frequently so-called *quantum wells* are used. They consist of two closely spaced heterointerfaces as shown in Fig. 17.3a. In this figure we plot the energy of the bottom of the conduction band as a func-

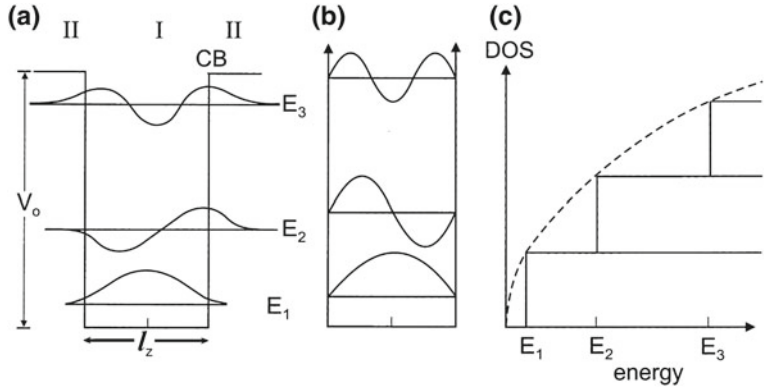


Fig. 17.3 The confining potential for electrons in a quantum well (**a**) and the idealized situation of infinitely high barriers (**b**) as well as the related density of states (**c**)

tion of space coordinate z perpendicular to the layer structure. The properties of the valence-band top will be discussed below.

The band discontinuities at the two interfaces form, in the simplest approximation, a one-dimensional square well potential of finite depth V_0 and width l_z. An idealization with infinitely high barriers is shown in Fig. 17.3b. We now define the potential at the bottom of the well as zero and the zero of the z-axis in the middle of the well. The solutions of this textbook problem with increasing energy are wavefunctions alternating between even and odd parity of the cosine- and sine-type, i.e., for even n_z

$$\Psi_n(\boldsymbol{r}) = \chi_n(z)\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}_\parallel \boldsymbol{r}_\parallel} = V^{-\frac{1}{2}}\, \mathrm{e}^{\mathrm{i}(k_x x + k_y y)} \cos\left(n_z \frac{\pi}{l_z} z\right) \tag{17.4a}$$

and an analogous expression with sine for odd n_z. The energies of the quantized electrons are given by

$$E_n(k_\parallel) = \frac{\hbar^2 \boldsymbol{k}_\parallel^2}{2m_\mathrm{e}} + E_{n_z} = \frac{\hbar^2 \boldsymbol{k}_\parallel^2}{2m_\mathrm{e}} + \frac{\hbar^2 \pi^2 n_z^2}{2m_\mathrm{e} l_z^2} = \frac{\hbar^2}{2m_\mathrm{e}} \left(\boldsymbol{k}_\parallel^2 + n_z^2 \frac{\pi^2}{l_z^2}\right) \tag{17.4b}$$

with $n_z = 1, 2, 3, \ldots$. This means there is an infinite number of quantized states E_{n_z} and at every E_{n_z} a new *subband* starts with constant density of states as shown in Fig. 17.3c. For the more realistic case of a finite potential V_0 (Fig. 17.3a) the wave functions penetrate exponentially decaying into the barrier.

Since the problem has inversion symmetry, the wave functions again have even or odd parity. From the Schrödinger equation we obtain for the envelope function in z direction:

$$-\frac{\hbar^2}{2}\frac{\partial}{\partial z}\left(\frac{1}{m_\mathrm{e}(z)}\frac{\partial}{\partial z}\right)\chi_{n_z}(z)+V(z)\,\chi_{n_z}(z)=E_{n_z}\,\chi(z) \tag{17.5a}$$

with

$$V(z)=\begin{cases}0 & \text{for} \quad |z|\le l_z/2\\ V_0 & \text{for} \quad |z|>l_z/2\,.\end{cases} \tag{17.5b}$$

The term $m_\mathrm{e}(z)$ allows to include different effective masses in well (material I) and barrier (material II).

The ansatz for the wave function $\chi_{n_z}(z)$ is

$$\chi_n(z)=\begin{cases}\pm B\exp(\kappa z) & \text{for} \quad z\le -l_z/2\\ A\begin{Bmatrix}\cos\\ \sin\end{Bmatrix}(kz) & \text{for} \quad -l_z/2\le z\le l_z/2\\ B\exp(-\kappa z) & \text{for} \quad z\ge l_z/2\end{cases} \tag{17.5c}$$

with

$$k=\sqrt{\frac{2m_\mathrm{e}^\mathrm{I}E_{n_z}}{\hbar^2}};\quad \kappa=\sqrt{\frac{2m_\mathrm{e}^\mathrm{II}\,(V_0-E)}{\hbar^2}}\,, \tag{17.5d}$$

which fulfills the necessity for normalizability. The boundary conditions at the interfaces require the steadiness of both the wave function $\chi(z)$ and the current $\frac{1}{m_\mathrm{e}}\frac{\partial}{\partial z}\chi(z)$ for stationary solutions. This leads to a transcendental equation, which can be solved graphically or numerically as detailed e.g., [02B1].

The main results are that there exists only a finite number of *quantized states* as shown schematically in Fig. 17.3a below V_0 and a *continuum* of extended states above V_0. The first few quantized states also follow, to a good approximation, a n_z^2 law, but with a slightly smaller prefactor than in (17.5b) due to the finite barrier height.

A short word on nomenclature is necessary at this point: The band gap of a given semiconductor (e.g., GaAs) is a property of the bulk material. The energetic distance between the first quantized electron and hole subbands in a type I quantum-well structure is then given by the sum of the band gap of the well material plus the two quantization energies $E_{\mathrm{e,h}}$. This sum should not be called band gap.

We have seen in Sect. 15.2 that the top of the valence band is usually more complex around the Γ-point than the conduction-band minimum. Even if we neglect the spin–orbit split-off band Γ_7 still the Γ_8 heavy-hole and light-hole bands are left in cubic semiconductors and the crystal field split A and B valence bands in wurtzite type materials. In first approximation we expect that the heavy-hole and light-hole subbands form two different series of subbands due to their different effective masses and the resulting difference in quantization energies as can be deduced, e.g., from (17.4). This is shown schematically in Fig. 17.4 where we plot the valence-band maximum for a quantum well and the envelope functions of the first four quantized levels.

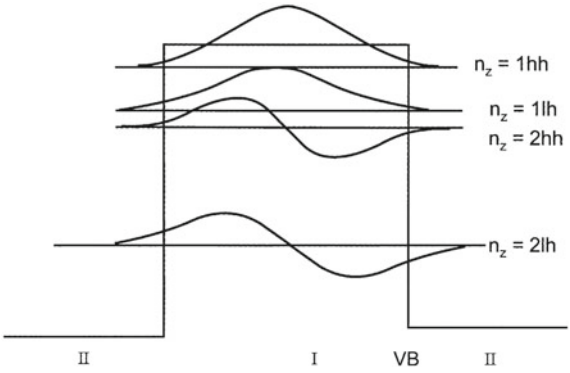


Fig. 17.4 The first two quantized levels for heavy and light holes (schematic)

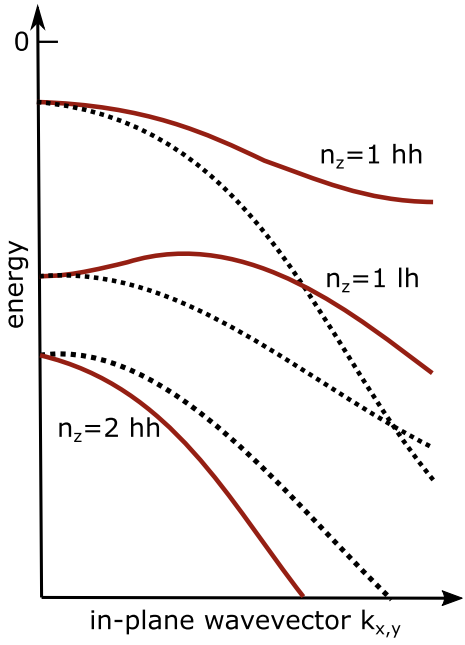


Fig. 17.5 Typical in-plane dispersion relation of the first three valence subbands in a GaAs/$Al_{1-x}Ga_xAs$ quantum well. The solid line includes the interaction of the subbands while the dotted line does not. For explicit calculations or experimental data see, e.g., [86B1, 91H1]

The next complication is that the wave function of the heavy hole in z-direction (which results in the quantization energies $E_{n_z\mathrm{hh}}$) exhibits a light-hole dispersion in the plane of the well and vice versa as was explained schematically for $p_{x,y,z}$ orbitals in Fig. 15.4 (instead of the real sp^3 hybrid wave functions). Consequently, the in-plane dispersion relations tend to cross for finite $\boldsymbol{k}_\parallel$. In fact, this crossing is avoided due to the coupling between the various subbands, resulting in a rather complex in-plane dispersion relation as shown in Fig. 17.5 for a GaAs quantum well embedded between $Al_{1-y}Ga_yAs$ barriers. As can be seen, the dispersion relation sometimes deviates strongly from a parabolic one and may in some intervals of $\boldsymbol{k}$ even obtain the opposite curvature. Consequently there will be deviations from the constant two-

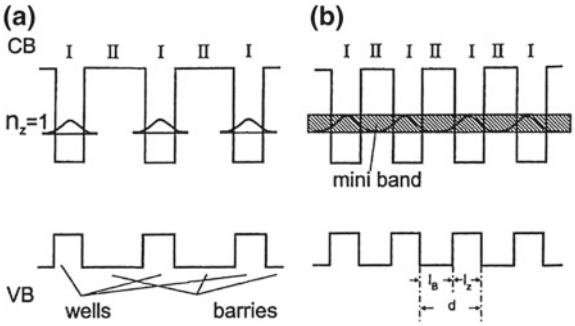


Fig. 17.6 The spatial variation of CB and VB extrema of a multiple quantum well (MQW) (**a**) and of a superlattice (SL) (**b**)

dimensional density of states expected in the effective-mass approximation. Due to these corrections, the 1hh dispersion in the plane fortunately shows a dispersion that comes close to m_{hh} while the 2hh states obtain an in-plane effective mass, which is even smaller than the light-hole mass.

The above consideration is valid if the quantum well is grown on a (100) plane. For high-index planes, things may get even more complicated as detailed in [09I1].

A crucial quantity for the calculation of the quantized states is the distribution of the total band-gap difference between well and barrier on the conduction (CB) and valence band (VB) offsets Δ_{CB} and Δ_{VB}, respectively. Even for a type I band alignment, where electrons and holes are confined in the same material (see Fig. 17.9) it is not trivial to determine the ratio $\Delta_{\mathrm{CB}}/\Delta_{\mathrm{VB}}$.This ratio may depend on the composition of the two materials, on the crystallographic orientation of the interface and on possible dipole layers at the interface. For the most common material combination for quantum wells $\mathrm{GaAs/Al_{1-x}Ga_xAs}$ with $x \geq 0.6$ i.e. for $\mathrm{Al_{1-x}Ga_xAs}$ still being a direct-gap semiconductor a value of $\Delta_{\mathrm{CB}}/\Delta_{\mathrm{VB}} \approx 0.60/0.40$ is accepted [89C1, 93A1]. For more data on the offsets in III–V quantum-well structures see [01K1], for II–VI structures see [04K1].

Until now we have treated only *single quantum wells* (SQW) and the situation for electrons and holes separately. Now we want to introduce the terms multiple quantum well, coupled quantum wells and superlattice and discuss the various types of band alignment.

In Fig. 17.6a we assume, that the band discontinuities of CB and VB are such that electrons and holes are confined in the same material. This is a so-called type I band alignment. If the width of the barrier is so large, that the wave functions of the first quantized electron states of adjacent wells do not overlap as in Fig. 17.6a we speak of a *multiple quantum well* (MQW). The density of states still increases as a sum over Heavyside functions at every E_{n_z} only the prefactor grows in proportion to the number of identical QW in the MQW.

If we now start to make the thickness of the barriers thinner and thinner we come to a regime, where the exponential tails of the wave functions in the barriers start to overlap as shown in Fig. 17.6b. If we couple only a few, e.g., two quantum wells in this way we obtain evidently *coupled quantum wells*. If they are identical, the ground

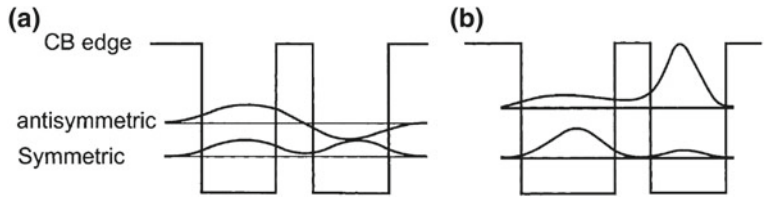


Fig. 17.7 Two coupled identical (**a**) and different (**b**) quantum wells and the envelope functions of the lowest states (schematic). In **a** the envelope wave function is given, to show clearly the symmetric and anti-symmetric superposition, in **b** only the modulus is shown

state splits into a symmetric and an anti-symmetric wave function, the latter being slightly higher in energy as shown in Fig. 17.7a. For coupled wells of different width as in Fig. 17.7b parity is no longer a good quantum number and the lowest states of the wide and the narrow well tend to obtain a slight admixture of the other wave function.

If we couple many identical quantum wells via sufficiently narrow barriers as shown in Fig. 17.6b we obtain in the z direction a new periodicity $d = l_z + l_b$ and the structure is known as a (one-dimensional) *superlattice (SL)*. Ideally d is given by a sum of integer multiple mono-layers of materials I and II.

As already discussed in Sect. 13.1 the additional periodicity in the superlattice results for the k_z direction in a *mini-Brillouin zone* extending from

$$-\pi/d \leq k_z \leq \pi/d \tag{17.6}$$

much smaller than the extension in the k_x and k_y direction, and a series of minibands with a finite curvature as shown schematically in Fig. 17.8a, b for $k_{\|} = 0$. These bands arise from the quantization levels E_{n_z} in the uncoupled (M)QW. The width of the minibands is typically a few meV and increases with increasing band index and decreasing barrier width since the coupling, which determines the bandwidth, increases in both cases (Fig. 17.8c).

The curvature and shape of the bands in the k_x, k_y plane remains essentially the one of the QW. In Fig. 17.8d, e we finally compare the density of states for a SL with the one of a MQW. Instead of a sum of Heavyside functions starting at E_{n_z} there is a steady variation of DOS(E), starting with a square-root behavior of a M_0 critical point and reaching the constant 2D density of states in a M_1 type behavior. Similar minibands can form for the various quantized hole states. If the hole mass is heavier, the penetration of the wave function into the barrier and the width of the minibands are smaller.

If the band discontinuities and the band alignment is such that electrons and holes are confined in the same material as shown, e.g., in Fig. 17.9a, we speak of a *type I band alignment*. The arrangement of the band edges as in Fig. 17.9b are called *type II staggered*. The electrons are confined in material I and the holes are essentially confined to material II. The transition is said to be spatially indirect. In Sect. 3.2 we

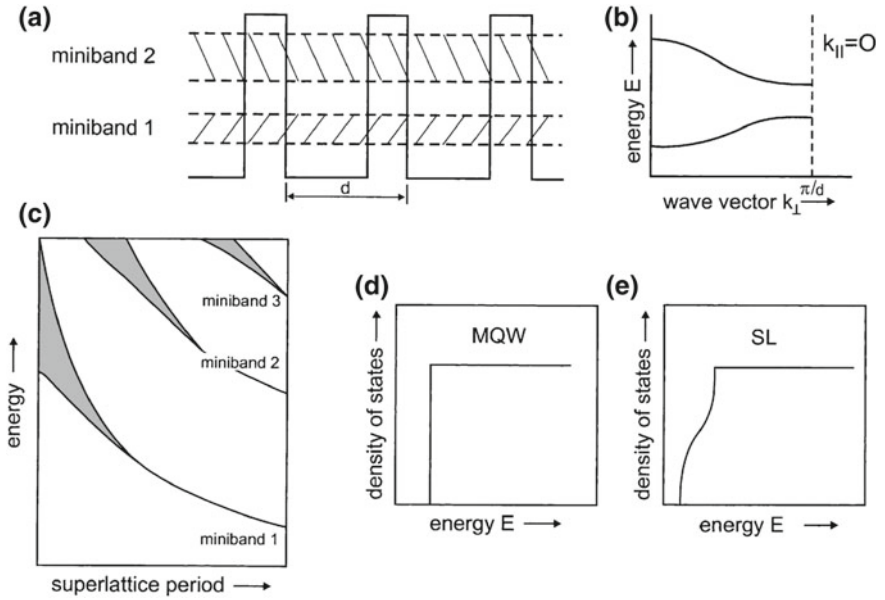


Fig. 17.8 Schematic summary of the properties of electronic bands in SLs: Minibands in the z direction ($k_x = k_y = 0$) of a SL (**a**, **b**), the width and position of these minibands as a function of the superlattice period (**c**), and the density of states (**e**) compared to the one of MQW (**d**)

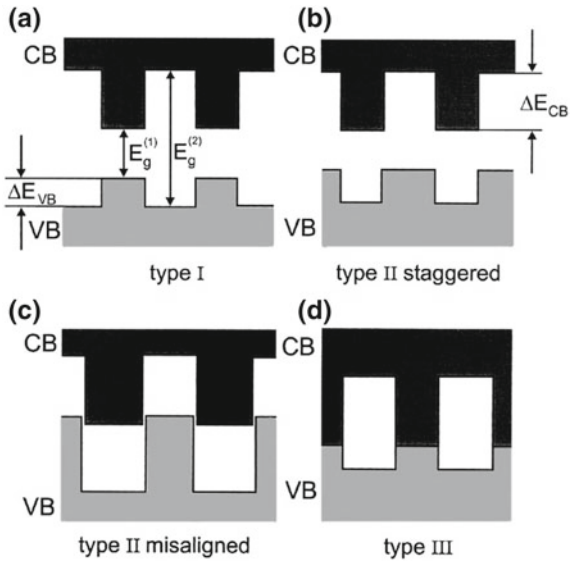


Fig. 17.9 Various band alignments in MQW or SL: type I (**a**), type II staggered (**b**), type II misaligned (**c**), type III (**d**)

will see an example for a type I → type II transition as a function of layer thickness in GaAs/AlAs superlattices. The arrangement of Fig. 17.9c where the bottom of the conduction band of material I falls below the top of the valence band of material II is known as *type II misaligned*. The combination of a semiconductor ($E_g > 0$) with a semimetal ($E_g \approx 0$) forms a *type III band alignment* structure (Fig. 17.9d). Examples for the various cases can be found, e.g., in [01K1] and in the next sections.

It should be mentioned that it is also possible to produce almost parabolic confinement potentials by a suitable and continuous spatial variation of the composition $x(z)$ in $Al_{1-x(z)}Ga_{x(z)}As$ from a maximum value (e.g., GaAs) towards a minimum. In this case the energetic spacing between the first few quantized levels E_{n_z} is almost constant as can be expected for a parabolic harmonic-oscillator like potential.

An almost parabolic confinement potential is also obtained in so-called *nipi structures* or *doping superlattices*. These structures do not contain abrupt or continuous hetero-junctions but are made from one material but with alternating n and p doping. Usually the differently doped layers are separated by undoped intrinsic layers explaining the name nipi structure. Electrons from the donors and the holes from the acceptors recombine, leaving the space-charge layers of D^+ and A^- behind. By solving the Poisson equation for the electrostatic potential $\varphi(\boldsymbol{r})$

$$\nabla^2\varphi(\boldsymbol{r}) = -\frac{\varphi(\boldsymbol{r})}{\varepsilon\varepsilon_0} \tag{17.7}$$

one obtains in growth direction a periodic modulation of the conduction and valence band edges as shown in Fig. 17.10.

The effective band gap of the spatially indirect transition indicated in Fig. 17.10 is much smaller than that of the parent material. The confinement potential is parabolic in the vicinity of the extrema resulting in energetically equidistant subband levels for electrons and holes.

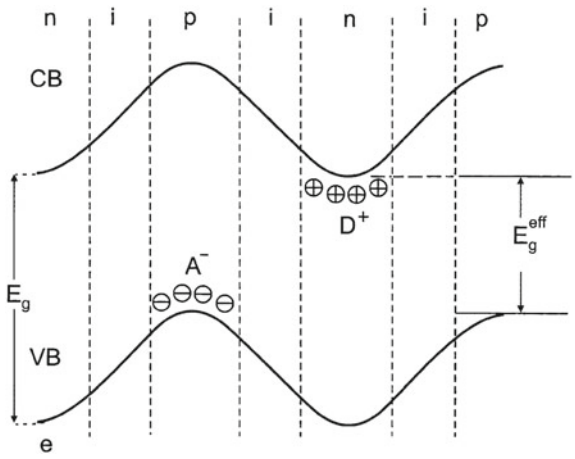


Fig. 17.10 The band alignment in a doping superlattice or nipi structure

The spatial overlap between the wave functions in the first confined states of electrons and holes is small. Consequently their lifetime is long. This makes such structures very sensitive, but slow photodetectors.

In some cases MQW structures are incorporated into nipi structures for special purposes. For more details see, e.g., [86D1, 92M1].

17.3 Monolayer Semiconductors

So far we have considered only "quasi-2D" semiconductor structures since growth of a perfect monolayer of a tetrahedrally coordinated material between two barrier layers is technically not feasible. This situation is completely different for the new class of 2D materials: they can form stable monolayers with strong in-plane binding and weak van der Waals–binding to adjacent (barrier) layers. Thus they show specific 2D behavior. Their electronic properties can easily be tuned by the choice of number of layers in well or barriers of hetero-structures, but also by the relative atomic alignment, by charge transfer or interaction between well and barriers (for reviews see, e.g., [16C1, 16M1, 16N1]).

The band-gaps of monolayer materials show a wide spread in energies and they can be insulators, semiconductors, semi-metals or metals [16C1]. Besides *graphen*, which has a zero bandgap, we will describe in the following mainly the properties of semiconducting materials with fundamental optical transitions in the visible and adjacent spectral regions. Most prominent representatives of this material's class are group-IV *transition-metal dichalcogenides (TMDs)* like MoS_2, WS_2 and $MoSe_2$, phosphorene, group-IV monochalcogenides like SnS, SnSe, and GeSe, In and Ga monochalcogenides like InSe, GaSe and GaS, and finally hexagonal boron nitride (hBN) [16C1, 16N1].

Fabrication of monolayer semiconductors is a difficult task. Mostly used is the method of micro-mechanical cleavage of monolayers from bulk material but also epitaxial methods are employed (see Appendix D.2 for details).

Most monolayer materials have a hexagonal Brillouin zone related to a triangular Bravais lattice in real space. The graphen *honeycomb lattice*, the resulting Brillouin zone and the relevant notation of basis vectors and high-symmetry points is given in Fig.17.11 [09C1]. The example of a TMD is given Fig. 17.12 [16M1]. In both cases the fundamental gap between valence and conduction band occurs at the K and K′ points of the Brillouin zone (actually in graphene this gap has zero energy, see below).

The K and K′ points are non-equivalent points of the Brillouin zone if the crystal inversion symmetry is broken. This is the case for the TMD lattices, but not for graphene where all sites of the honeycomb are occupied by the same type of atoms. However, such symmetry breaking can be induced in graphene by application of an external electric field [17D1]. Spin-orbit coupling now leads to major consequences: first there are a spin-orbit splitting of the conduction bands and a much larger one of the valence bands. The spin degeneracy at each valley is thus lifted but the spin-

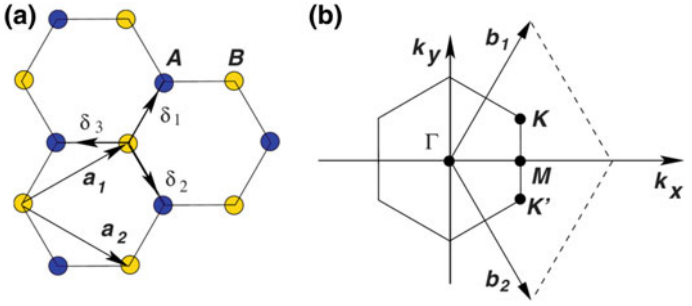


Fig. 17.11 Honeycomb lattice structure of graphene, made out of two interpenetrating triangular lattices (a1 and a2 are the lattice unit vectors, and δ_i, i = 1, 2, 3 are the nearest-neighbor vectors) (**a**); corresponding Brillouin zone (**b**) (reprinted by permission from [09C1] ©2009 American Physical Society https://doi.org/10.1103/RevModPhys.81.109)

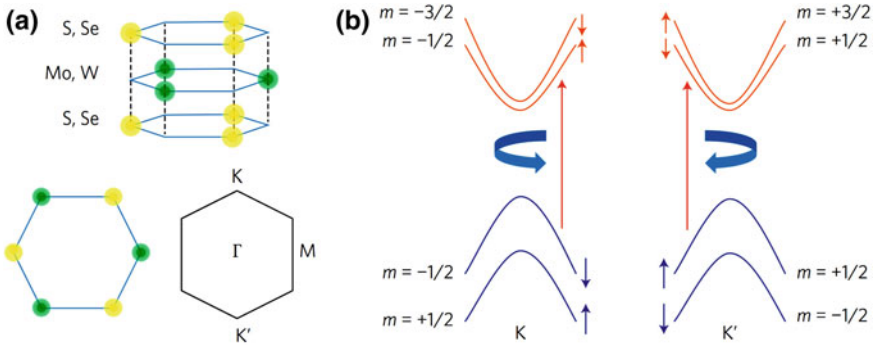


Fig. 17.12 Illustration of the hexagonal lattice and Brillouin zone of WX_2 (**a**). Sketch of the valley-dependent optical selection rules at the band edges. In the K (K′) valley, the optical transitions couple to σ^+ (σ^-) polarized light according to the selection rule $\Delta m = \pm 1$ with m being the projection of the total angular momentum; see Fig. 23.19 for a scheme explicitly including excitonic effects (**b**) (reprinted by permission from [16M1] ©2016 Springer Nature)

up state at the K valley stays degenerate with the spin-down state at the K′ valley. The spin-orbit interaction actually couples of the spin and valley degrees of freedom (the *valley degree of freedom* can be treated theoretically as a pseudo-spin). The consequence of this so-called *spin-valley locking* is, that the occupation of the valleys at K or K′ can be controlled, e.g., optically by choice of the respective circular polarization of absorbed photons (see Fig. 17.12) [12M1, 12X1].

The spin-valley locking enables the field of *valleytronics* since the electronic transport properties are valley dependent. Even non-dissipative pure valley currents without charge transport (similar to the case of pure spin currents) are observed (for reviews see, e.g., [15S1, 16M1]).

We now want to focus on some specific properties of monolayer materials. We start with the TMDs [12W1]. The electronic bandstructure of TMDs is characterized

by the dominant contribution of d-electron bands to both conduction and valence bands. The band extrema are located at the K and K′ corners of the Brillouin zone as has been discussed above. Some TMDs like $MoSe_2$ and WS_2 have an indirect bandgap for the bulk material (and even for bilayers) but acquire a direct gap due to confinement in the monolayer and change in the hybridization of the electron wavefunctions [11K1].

Actually, the optical transitions described in connection with spin-valley locking in TMDs are related to strongly bound excitonic species. The large binding results from large effective masses of electrons and holes, reduced screening typical for (quasi-)2D systems, and carrier confinement in a single monolayer [15W1, 16M1]. Please note that the schematic depiction of the optical transitions of Fig.17.12, which is used in a similar form in most publications, is misleading or even wrong when discussing excitons. The exciton is a composite particle with a dispersion relation describing its center-of-mass motion. Although the properties of the electron and the hole derive from the single-particle approximation used in bandstructure scheme, the exciton cannot be drawn in the same figure together with the electron bands. Care has to be taken when discussing selection rules for excitonic transitions. This discussion gets quite awkward when the one-electron bandstructure is used rather that the exciton states, in particular for dark excitons. We will return to these excitonic properties in Sect. 23.3 and show a better schematic depiction of the free-carrier and excitonic dispersions in Fig. 23.19.

The best studied monolayer material with highly promising application in (opto-) electronics is graphene [07G1, 10B1] where the 2D arrangement of the carbon atoms is due to the sp^2 hybridization. But, pure layers of graphene are not semiconducting since they have no band-gap which limits their applicability in electronic devices [12N1]. Valence and conduction bands touch each other at the K-points of the Brillouin zone and have linear dispersion relations [07C1, 09C1, 10B1] (see Fig. 17.13). The electrons behave like mass-less particles described by the relativistic Dirac rather than the Schrödinger equation. This results in extraordinary electronic transport properties in the quantum regime [07G1, 09C1]. Quasi-ballistic electron transport at room temperature combined with high mechanical stability and flexibility makes graphene a rising star in nano-electronics.

Graphene shows a strong light-matter interaction due to unusual band structure. A single layer of graphene is of course transparent (it absorbes 2.3% of incident light) but its transmittance $T \approx (1 - \pi\alpha)$ is directly given by the fine-structure constant α [08N1]. This relationship to a natural constant results from the 2D character of the electrons and the gap-less bandstructure. Due to the linear dispersion absorbance of graphene layers there is an optical transition available for each wavelength of light resulting in broad-band absorption. Due to Pauli blocking the absorption can easily be saturated which in combination with fast carrier relaxation can be used for high-speed optical modulators. Transparency and efficient carrier transport make graphene suitable as transparent electrodes in various devices like solar cells [10B1].

Luminescence from graphene is intrinsically inefficient due the missing band-gap and due the fast carrier relaxation. A gap can in principal be induced by further reducing dimensionality to nano-ribbons or quantum dots. Bright luminescence is,

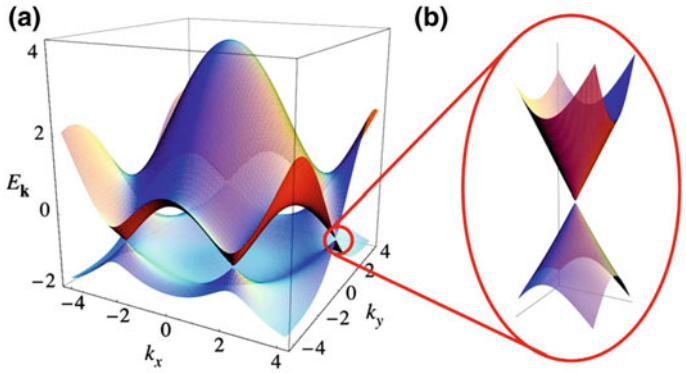


Fig. 17.13 Electronic dispersion of graphene with the energy given in units of the nearest-neighbor hopping energy $t = 2.7$ eV (**a**); close-up of the energy bands close to one of the Dirac points (**b**) (reprinted by permission from [09C1] ©2009 American Physical Society https://doi.org/10.1103/RevModPhys.81.109)

however, better achieved by introducing defects by oxidation. Light emission in the whole spectral range of interest has been achieved in such oxidized graphen as part of hybrid hetero-structures [10B1].

Similar to graphen, but with two different atoms in the hexagonal lattice resultuing from sp^2 hybridization, is hBN. It has a large bandgap in the UV and is often used as barrier layer in heterostructures with graphen or TMDs. The strong anisotropy between in-plane and out-of plane permittivity leads to strongly confined phonon–polariton modes in the mid-infrared with hyperbolic dispersion relation [16C1, 16N1].

Phosphorene is a mono-elemental monolayer semiconductor with an orthorombic lattice. This results in an in-plane anisotropy of the optical properties. The same holds for group-IV mono-chalcogenides. In the latter inversion symmetry is broken resulting in sizable spin-orbit splitting and piezoelectric effects. They have mainly indirect bandgaps and two pairs of twofold degenerate valence and conduction bands situated at main axes in the othorombic Brillouin zone. The valley pairs can be addressed by linearly polarized light [16N1].

Regarding possible applications of monolayer materials one can summarize the following. The flat dispersion of the electronic bands typical for many monolayer semiconductors leads to a large density of states and thus high absorption which is beneficial for photo-detectors. P-n junctions can be formed with possible applications as solar cells or light emitting diodes. Here hetero-structures of monolayer materials are useful to engineer the band gap. Graphen can be used in hetero-structures as electrodes to extract carriers. Large-gap hBN is useful as barrier material to increase the confinement of carriers, e.g., in TMDs [16N1].

The optical properties of semiconducting monolayer materials are strongly influenced by excitonic features. We will discuss these excitonic effects in Sect. 23.3.

17.4 Quantum Wires and Nanorods

If the motion of the carriers are confined in two directions, the z and y directions, by a suitable potential, e.g., a discontinuity of the band gap in a hetero-structure or the transition to vacuum or air, there remains only free motion in one direction. Consequently, such structures are called quasi-one dimensional or *quantum wires*.

The energies of the carriers are given in an effective-mass approximation by

$$E(k_x) = E_{n_i,n_j} + \frac{\hbar^2 k_x^2}{2m_{\mathrm{e,h}}} \tag{17.8}$$

where we assume that the only remaining direction of free motion is along the x-axis. At every energy level E_{n_i,n_j} starts a quasi-one-dimensional subband. The singularity of the density of states according to Fig. 17.1 expected at the edge of each energy subband is more or less smeared out by wire-width fluctuations, inter-diffusion or segregation. This is similar to the case of quantum wells as discussed in Appendix D.1.

The quantization energy E_{n_i,n_j} depends on the cross section of the wire [02K2]. This cross section, however, may be rather complex depending on the realization of the wire. The growth of quantum wires is detailed in Appendix D.3.

Quantum wires which are lithographically defined from QWs as shown in Fig. 17.14a have a nearly rectangular cross section. The electronic confinement is given by the barriers of the QW (E_{n_z}) and in the y-direction by the transition to vacuum (air) or to an overgrown barrier material. Here, the confinement energies are simply additive, i.e.,

$$E_{n_i,n_j} = E_{n_z} + E_{n_y}. \tag{17.9}$$

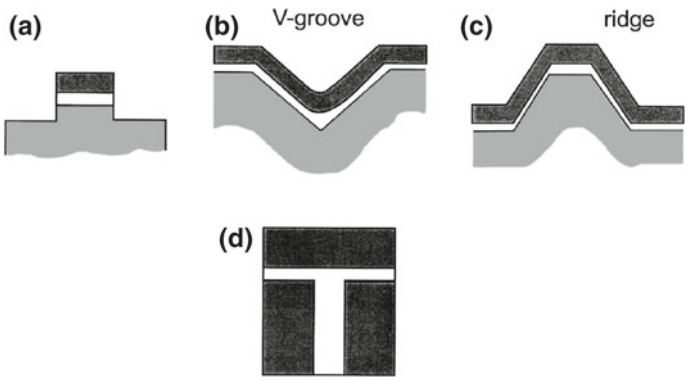


Fig. 17.14 Various possibilities to create potentials forming quasi-one-dimensional quantum wires. Lithography and etching of quantum wells (**a**), growth in selectively etched V-grooves (**b**) or on ridges (**c**), and T-shaped quantum wires occurring at the junction of two orthogonal quantum wells. In all cases the wire extends normal to the paper plane

A problem in these structures is that the width of the quantum well l_z typically has values between 3 and 10 nm while the best values, which can be reached by lithography and dry or wet etching are usually in the range of 20 nm $\leq l_y \leq$ 50 nm. Since the quantization energy varies roughly with l_i^{-2} one finds that usually $E_{n_z} \gg E_{n_y}$ for $n_z = n_y$. Often it is already considered as an argument for confinement in two dimensions if, e.g., the excitonic transition energies are blue-shifted by a few meV compared to the quantum well (see examples in Sect. 23.4).

To achieve better controlled confinement many alternatives to form quantum wires have been established. We show three of them in Fig. 17.14b–d, namely V-groove, ridge and T-shaped QWRs. In the first two cases a quantum well grown on a pre-patterned substrate is thicker on the bottom of the V-groove or on the top surface of the ridge than on the side flanks of the structures or on the flat parts of the substrate. Since the confinement energy increases quadratically with decreasing well width this effect produces a confinement potential along the curved quantum well in addition to the potential step normal to it. Both confining potentials produce together the confinement in two, here usually not orthogonal, directions and leaving a free motion of the carriers only in x-direction. For more details see, e.g., [89K1, 99B1, 00B1, 02K1].

So-called *cleaved-edge overgrowth* (Fig. 17.14d) leads to a region, where two quantum wells meet. The "average width" of the confining potential is wider than in each of the two quantum wells. This again produces a confining potential in two directions [85C1, 96L1, 97H1, 98A1]. If one starts from a multiple quantum well or a superlattice one obtains by the cleaved-edge overgrowth an array of uncoupled or coupled quantum wires, respectively.

Quantum wires of CdSe with small diameters of few nanometer can be grown (like quantum dots) from solution to form colloidal nanostructures [02K2, 03M1]. Quantum efficiency of optical emission can be enhanced in such wires when core/shell structures are grown. i.e., when the CdSe wire is enclosed with a large band-gap material like ZnS [03M1]. These wire have a columnar structure and allow to investigate the dependence of the electronic level structure on wire dimensions. Blue shifts of up to about 200 nm can be observed [02K2].

The reduced dimension and symmetry of the quantum wires leads besides the mentioned quantization effects also to band mixing. Analytical calculations of quantum-wire band structures and the resulting optical matrix elements are found in [90S1, 91S1]. The non-centrosymmetric crystallographic lattice of CdSe wires leads to a strong dipole moment oriented along the c-axis [03L2].

To conclude this section we want to focus on thin columnar crystals, so-called *nanorods*, which grow on substrates in a self-organized fashion. Typical dimensions are several μm length and a few tens or hundreds of nm thickness.

Prominent materials for nanorods are ZnO or GaN. The confinement energy of these rods perpendicular to their growth direction is usually negligible for carriers and excitons, but they form efficient one-dimensional wave guides due to the large index step between semiconductor and vacuum (or air). Partly quantum wells or superlattices are incorporated into these rods in radial or longitudinal direction. For

some reviews see [01H1, 01H2, 01P1, 03L1, 03Z1, 04T1, 04Z1, 07K1, 09W1, 10K1, 10K2].

Quantum wires or nano-wires/rods offer tunability of the optical properties and are considered as bridging the gap between macro- and nano-sized elements in nano-photonics due to their large aspect ratio. Large efforts are taken to integrate them in electronic and opto-electronic devices like transistors, displays, solar cells, batteries, sensors, and light emitting devices. For a review on nano-wire photonics see, e.g., [09Y1], for nano-wire lasers see [12Z1], for nano-wire solar cells see [16M2] for sensors see [01C1], for electrical properties of Si nano-wires see [09S1].

17.5 Quantum Dots

The ultimate limit in reducing the quasi-dimension is confinement in all three directions of space leaving no more possibility for free propagation of the carriers. The resulting quasi-zero dimensional structures are mainly termed *quantum dots*, but also quantum boxes, and nano-crystals are common names.

The energy levels in a single quantum dot are discrete (see Fig. 17.1), i.e.,

$$E = E_{n_i,n_j,n_k} \tag{17.10}$$

like in atoms below the ionization threshold. This is the reason why quantum dots are sometimes called *artificial atoms*. We will discuss the influence of the quasi-0D confinement on the electronic states in detail in Sect. 23.5. Since the size of the quantum dots and the excitonic Bohr radius are similar one has to take care of the Coulomb interaction when calculating electronic or excitonic energies.

Discrete electronic states are observed in spectroscopy of quantum dots (Sect. 23.5) if the experiments are performed on individual dots which have to be isolated by techniques with high spatial resolution (see 23.5.2). Typically a sample contains an ensemble of many quantum dots with unavoidable size fluctuations between the individuals. This occurs independently of how these quantum dots have been formed. Since the quantization energy depends on the dot size, this gives one contribution to the inhomogeneous broadening of the ideal δ-function density of states for the energy spectrum of the ensemble.

Quantum dots actually have a variety of manifestations derived from principally different fabrication routes [95W1]. Most important are colloidal QDs forming by pyrolysis in hot liquids [93M1, 96A1], self-assembled quantum dots (SAQD) established in strained-layer overgrowth [99B2], QDs grown on pre-structured substrates [11H1], quasi-0D confining potentials realized by interface fluctuations of quantum wells [96G1], and electrostatically defined QDs [97K1]. The respective fabrication techniques, except for the last case, are reviewed in Appendix D.4.

Colloidal QDs have a very high photoluminescence yield and are thus used in LED displays or as optical markers in microscopy of biological samples. They are often produced in the for of core/shell quantum dots. These are dots which have a

core of one material covered by a shell of a second one. The material combinations can be of type I or II concerning the band alignments. If in the first case the shell has the larger band gap (as is the case e.g., for CdS/ZnS core/shell structures), this arrangement allows to protect the core e.g., against surface effects. In contrast the combinations CdSe/CdTe or CdSe/ZnTe result in type-II band alignments. For a review on such dots and their applications see, e.g., [13S1].

Self-assembled quantum dots (SAQDs) can be fabricated in strongly lattice mismatched systems using the *Stranski-Krastanov growth mode*. After overgrowth of the hillocks or pyramids by a barrier layer one finds quasi-0D confinement in island-shaped QDs typically embedded in a wetting layer. The preferential systems in which these SAQDs or quantum islands are investigated are $In_{1-y}Ga_yAs/GaAs$ including $y = 0$ [99B2] and CdSe/ZnSe [04S1].

The SAQDs have often stronger confinement in growth direction compared to the one in-plane. The actual chemical composition of the QD is typically affected by intermixing of the components. A lot of effort has to be invested to calculate the electron and hole eigenstates and energies in these structures including the spatially varying strain [99B2].

The AlAs/GaAs system is not prone to significant inter-diffusion and segregation. But, thickness fluctuations occur readily in GaAs quantum wells as schematically shown in Fig. D.3. These fluctuations can lead to 0D carrier confinement so that the term *interfacial quantum dots (IQD)* has been introduced in the literature. The 0D-character of the optical transitions can directly be deduced from the appearance of spectrally sharp emission lines in spatially resolved spectroscopy (see Sect. 23.5 and, e.g., [96G1, 00L1, 03M2]).

Finally we want to address shortly *electrostatically defined QDs* [97K1]. These dots are realized in form of quasi-0D potential minima in a 2D electron gas. These potential minima are induced by applying suitable electrostatic potentials to micro-electrode structures deposited on the barrier material over a quantum well. They are important elements of nano-electronic devices exploiting transport properties like Coulomb blockade, and single-electron tunneling or electron turn style. They are also used to realize spin Qubits [16T1]. Since optical transitions in these systems fall in the micro-wave regime, we will not further consider them here.

Problems

17.1 Calculate the effective density of states per unit volume at 300 K for electrons and holes in bulk GaAs and the density of states per unit area in a GaAs QW. Calculate the corresponding density per volume for $l_z = 10\,\text{nm}$.

17.2 Calculate the positions of the first three quantized electron levels for a GaAs quantum well assuming infinitely high barriers and the realistic band discontinuity to $Al_{0.40}Ga_{0.60}As$ and compare them.

References

[72W1] F. Wooten, *Optical Properties of Solids* (Academic, New York, 1972)
[85C1] Y.-C. Chang, L.L. Chang, L. Esaki, Appl. Phys. Lett. **47**, 1324 (1985)
[86B1] G. Bastard, J.A. Brum, IEEE J. QE- **22**, 1625 (1986)
[86D1] G.H. Döhler, IEEE J. QE- **22**, 1682 (1986)
[89C1] R.M. Cohen, Z.M. Fang, J. Appl. Phys. **65**, 612 (1989)
[89K1] E. Kapon, D.M. Hwang, R. Bhat, Phys. Rev. Lett. **63**, 430 (1989)
[90S1] P.C. Sercel, K.J. Vahala, Phys. Rev. B **42**, 3690 (1990)
[91H1] R.K. Hayden et al., Phys. Rev. Lett. **66**, 1749 (1991)
[91S1] P.C. Sercel, K.J. Vahala, Phys. Rev. B **44**, 5681 (1991)
[92M1] S. Malzer et al., Phys. Status Solidi (b) **173**, 459 (1992)
[93A1] M. Missous, in *Properties of Aluminium Gallium Arsenide*, EMIS Datareviews Series No. 7, ed. by S. Adachi (1993), p. 73
[93M1] C.B. Murray, D.J. Norris, M.G. Bawendi, J. Am. Chem. Soc. **115**, 8706 (1993)
[95W1] U. Woggon, S.V. Gaponenko, Phys. Status Solidi (b) **189**, 285 (1995)
[96A1] A.P. Alivisatos, Science **271**, 933 (1996)
[96G1] D. Gammon, E.S. Snow, B.V. Shanabrook, D.S. Katzer, D. Park, Phys. Rev. Lett. **76**, 3005 (1996)
[96L1] W. Langbein, H. Gislason, J. Hvam, Phys. Rev. B **54**, 14595 (1996)
[97H1] J. Hasen et al., Nature **390**, 54 (1997)
[97K1] L.P. Kouwenhoven et al., in *Proceedings of the NATO Advanced Study Institute on Mesoscopic Electron Transport*. Kluwer Series, E345 (1997), p. 105
[98A1] H. Akiyama, J. Phys. Condens. Matter **10**, 3095 (1998)
[99B1] G. Biasiol, E. Kapon, J. Crystal Growth **201/202**, 62 (1999)
[99B2] D. Bimberg, N. Grundmann, N.N. Ledentsov, *Quantum Dot Heterostructures* (Wiley, Chicester, 1999)
[00B1] G. Biasiol, K. Leifer, E. Kapon, Phys. Rev. B **61**, 7223 (2000)
[00L1] D. Lüerßen, R. Bleher, H. Kalt, Phys. Rev. B **61**, 15812 (2000)
[01C1] Y. Cui, Q. Wei, H. Park, C.M. Lieber, Science **293**, 1289 (2001)
[01H1] Q. Huang et al., J. Crystal Growth **227/228**, 117 (2001)
[01H2] M.H. Huang et al., Science **292**, 1897 (2001)
[01K1] C. Klingshirn in Landolt-Börnstein III/34: *Semiconductor Quantum Structures*, Subvol. C *Optical Properties*, Part 1, (Springer, Berlin, 2001)
[01P1] Z.W. Pan, Z.R. Dai, Z.L. Wang, Science **291**, 1947 (2001)
[02B1] P.K. Basu, *Theory of Optical Processes in Semiconductors: Bulk and Microstructures* (Oxford University Press, Oxford, 2002)
[02K1] E. Kapon, in *Spectroscopy of Systems with Spatially Confined Structures* (2001). NATO Science Series II, vol. 90 (Kluwer, Dordrecht, 2002), p. 243
[02K2] D. Katz et al., Phys. Rev. Lett. **89**, 086801 (2002)
[03L1] J.Y. Lao et al., Nano Lett. **3**, 235 (2003)
[03L2] L. Li, A.P. Alivisatos, Phys. Rev. Lett. **90**, 097402 (2003)
[03M1] T. Mokari, U. Banin, Chem. Mater. **15**, 3955 (2003)
[03M2] K. Matsuda et al., Phys. Rev. Lett. **91**, 177401 (2003)
[03Z1] B.P. Zhang et al., Appl. Phys. Lett. **83**, 1635 (2003)
[04K1] H. Kalt in Landolt-Börnstein III/34: *Semiconductor Quantum Structures*, Subvol. C *Optical Properties*, Part 2 (Springer, Berlin, 2004)
[04S1] V.A. Scubin, N.N. Ledentsov, D. Bimberg, *Epitaxy of Nanostructures* (Springer, Heidelberg, 2004)
[04T1] T.Y. Tan, N. Li, U. Gösele, Appl. Phys. A **78**, 519 (2004)
[04Z1] B.P. Zhang et al., Appl. Phys. Lett. **84**, 4098 (2004) ; J. Appl. Phys. **96**, 340 (2004)
[07C1] J.C. Charlier, P.C. Eklund, J. Zhu, A.C. Ferrari, Electron and phonon properties of graphene: their relationship with carbon nanotubes, in *Carbon Nanotubes*, Topics in Applied Physics, vol. 111 ed. by A. Jorio, G. Dresselhaus, M.S. Dresselhaus (Springer, Berlin, 2007)

[07G1] A.K. Geim, K.S. Novoselov, Nat. Mater. **6**, 183 (2007)
[07K1] C. Klingshirn, Phys. Status Solidi B **244**, 3027 (2007)
[08N1] R.R. Nair et al., Science **320**, 1308 (2008)
[09C1] A.H. Castro Neto et al., Rev. Mod. Phys. **81**, 109 (2009)
[09I1] H. Ibach, H. Löth, *Solid State Physics* (Springer, Berlin, 2009)
[09S1] V. Schmidt, J.V. Wittemann, S. Senz, U. Gösele, Adv. Mater. **21**, 2681 (2009)
[09W1] M. Willander et al., Nanotechnology **20**, 332001 (2009)
[09Y1] R. Yan, D. Gargas, P. Yang, Nat. Photonics **3**, 569 (2009)
[10B1] F. Bonaccorso, Z. Sun, T. Hasan, A.C. Ferrari, Nat. Photonics **4**, 611 (2010)
[10K1] C. Klingshirn et al., Phys. Status Solidi B **247**, 1424 (2010)
[10K2] C. Klingshirn, B.K. Meyer, A. Waag, A. Hoffmann, J. Geurts, *Zinc Oxide: From Fundamental Properties Towards Novel Applications*, Springer Series in Materials Science, vol. 120 (Springer, Berlin, 2010)
[10Y1] P.Y. Yu, M. Cardona, *Fundamentals of Semiconductors*, 4th edn. (Springer, Heidelberg, 2010)
[11H1] M. Helfrich et al., J. Cryst. Growth **323**, 187 (2011)
[11K1] A. Kuc, N. Zibouche, T. Heine, Phys. Rev. B **83**, 245213 (2011)
[12M1] K.F. Mak, K. He, J. Shan, T.F. Heinz, Nat. Nanotech. **7**, 494 (2012)
[12N1] K.S. Novoselov et al., Nature **490**, 192 (2012)
[12W1] Q.H. Wang et al., Nat. Nanotechnol. **7**, 699 (2012)
[12X1] D. Xiao et al., Phys. Rev. Lett. **108**, 196802 (2012)
[12Z1] C.Z. Ning, Semiconductor nanowire lasers, in *Advances in Semiconductor Lasers, Semiconductors and Semimetals*, vol. 86, 455, ed. by J.J. Coleman, A.C. Bryce, C. Jagadish (Academic, Elsevier, Burlington, 2012)
[13S1] Y. Shirasaki, G.J. Supran, M.G. Bawendi, V. Bulovic, Nat. Photonics **7**, 13 (2013)
[15S1] Y. Shimazaki et al., Nat. Phys. **11**, 1032 (2015)
[15W1] F. Wu, F. Qu, A.H. MacDonald, Phys. Rev. B **91**, 075310 (2015)
[16C1] A. Castellanos-Gomez, Nat. Photonics **10**, 202 (2016)
[16M1] K.F. Mak, J. Shan, Nat. Photonics **10**, 216 (2016)
[16M2] S. Mokkapati, C. Jagadish, Opt. Express **24**, 17345 (2016)
[16N1] K.S. Novoselov et al., Science **353**, aac9439 (2016)
[16T1] S. Tarucha et al., Spin qubits with semiconductor quantum dots, in *Principles and Methods of Quantum Information Technologies*, Lecture Notes in Physics, vol. 911 ed. by Y. Yamamoto, K. Semba (Springer, Tokyo, 2016)
[17D1] S. Dufferwiel et al., Nat. Photonics **11**, 497 (2017)

Chapter 18
Electronic Defects and Disorder

Abstract Deviations from the perfect crystalline periodicity in the form of defects and disorder play an important role in semiconductor optics. Donors and acceptors not only provide as dopants the free carriers in the bands. They also provide states for radiative recombination with photon energies below the band-gap energy. Such transitions are thus prominent in particular in luminescence spectra. This now allows for optical determination of the defect's characteristic properties. We will describe the wavefunctions and statistics of carriers bound to shallow and deep defects centers as well as typical recombination mechanisms including the ones involving additional phonons. Then we introduce the method of modulation doping in quantum wells. The second section reviews different types of localization mechanisms for carriers. Such localization is found in disordered potentials, e.g., for high levels of doping, in alloy semiconductors, in quantum wells due to well-width fluctuations or in amorphous semiconductors. We elaborate the modifications of the electronic density of states, e.g., in the form of band tails and show the respective consequences on optical spectra. This also includes the identification of regimes with different extend of the wavefunctions like the Urbach tail and the Tauc regime which are separated by the optical or mobility gap. We will close with the concepts of weak localization and percolation.

18.1 Defects, Defect States and Doping

In context with local phonon modes, we have already mentioned the imperfections which are present in all real crystals. Here we outline the electronic states connected with defects and concentrate again on point defect neglecting thus one dimensional defects like dislocations, two dimensional ones like stacking faults, grain boundaries, surfaces and interfaces, or three dimensional ones like precipitates and voids.

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_18

18.1.1 Donors and Acceptors in Bulk Semiconductors

Point defects are present, even in good materials, with densities of up to $10^{13} - 10^{17}\text{cm}^{-3}$. Values of $10^{9} - 10^{13}\text{cm}^{-3}$ which are given in literature, usually refer only to "electrically active" defects. This means that there is a huge difference between the concentrations of ionized donors and acceptors (see below) and defects, which are electrically not active but which may act still as scattering or recombination centers. The density of dislocations ranges from 0 to some 10^{4} cm^{-2} in high quality materials.

In Fig. 13.2 we classified the point defects according to the way in which they are incorporated in the lattice, e.g., as *interstitial* or *substitutional*. Now we consider their electronic properties and present *donors*, *acceptors*, *isoelectronic traps*, and *recombination centers* in Fig. 18.1a using the representation of the bands and defects over a real space coordinate. A donor is preferentially a shallow center which has an energy level just below the conduction band and can easily give an electron, e.g., by thermal ionization, to this band

$$D^0 \leftrightarrows D^+ + \mathrm{e}\,. \tag{18.1a}$$

In Fig. 18.1b we chose for a neutral acceptor (see below with (18.1b)) a representation in $\boldsymbol{k}$-space. The bound and spatially localized states have a finite extension in real space. Consequently $\boldsymbol{k}$-conservation is relaxed and so-called *free-to-bound transitions* (e.g., of a free electron with a bound hole as shown in Fig. 18.1a,b) become possible without $\boldsymbol{k}$-conservation. These transitions monitor directly the distribution of the free carriers in their band (see [04A1]). Please note that figures in which the donor and/or acceptor levels are shown in $\boldsymbol{k}$-space with a curvature identical to the one of conduction or valence band are wrong. See also the discussion with (18.3) below.

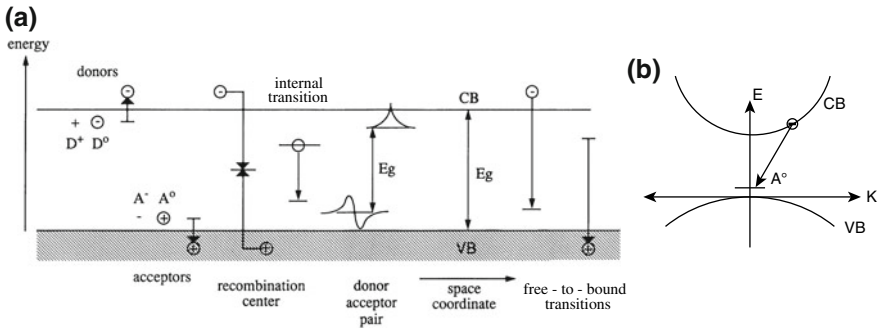


Fig. 18.1 Schematic drawing of various impurity levels in semiconductors using the presentation of the band structure in real space (**a**) and the recombination of a free electron with the bound hole of a neutral acceptor in an $E(\boldsymbol{k})$ diagram (**b**)

Donors are often formed by substitutional atoms situated in the periodic table one column to the right of the atom which they replace. Examples are N or P in Ge or Si, Si on Ga site in GaAs, or Cl on Se sites in ZnSe, Ga or In on Zn site in ZnSe etc. Furthermore, donors can be formed by interstitials which have a weakly bound electron such as H, Li or Na in some II–VI compounds (see Fig. 5.1 of [10K2]). Furthermore there are deep donors (and acceptor) situated energetically further away from the respective band edges.

In analogy, acceptors can easily accommodate an electron from the valence band, i.e., they emit a hole into it

$$A^0 \leftrightarrows A^- + \mathrm{h}. \tag{18.1b}$$

Acceptors may be formed by substitutional atoms which have one electron less than the one which they replace. Thus these impurities are often found in the periodic table to the left of the atom which they replace, for example, Ga or B in Si and Ge, Li or Na on the cation site and N on the anion site in II–VI compounds, or Si on the As or Be on Ga site in GaAs. Also deep acceptors are known. Especially in wide-gap semiconductors on encounters frequently the fact that they can be easily doped one type (e.g., n-type) and not or difficultly the other. Examples are e.g., GaN, ZnSe or ZnO, which are easily doped n-type beyond $n = 10^{20}\,\mathrm{cm}^{-3}$ but were or still are difficult to dope p-type in a stable and reproducible way. Especially in ZnO p-type doping is the prominent problem for wide applications in blue/UV optoelectronics. On the other hand ZnTe, Cu_2O or the Cu-halides tend to be p-type. For more details and some reviews see [95F1, 00N1, 02H1, 07K1, 08L1, 09L1, 09M1, 10K1, 10K2].

It is clear from some of the above-mentioned examples that the same atom can act as donor or acceptor depending on the way it is introduced into the lattice. This is one possibility for *self-compensation*. Another arises from the fact that some vacancies or interstitials can act as donors or as acceptors or that the incorporation of an acceptor (donor) in one of the II–VI compounds, which are notoriously n-type (p-type) like ZnSe, CdS or ZnO (ZnTe) triggers in thermodynamic equilibrium rather the formation of another defect, which acts as a donor (acceptor) instead of shifting the Fermi-level close to the valence (conduction) band to produce p-type (n-type) conductivity [88W1, 95F1].

Actually the success of p-doping of ZnSe is based entirely on freezing in a non-equilibrium substitutional N occupation on Se sites created during growth. A material which has a higher concentration of donors (acceptors) is called *n-type* (*p-type*) and we recall that $np = n_i^2(T)$ in thermodynamic equilibrium. The ability to choose the type and concentration of carriers over a wide range by doping with donors and/or acceptors is the basis for the widespread and important application of semiconductors in electronic devices like diodes and transistors. This topic is beyond the scope of this book and we refer the reader to text books on the physics of semiconductor devices, like [06S1, 16G1].

A shallow donor (acceptor) can be considered as a positively (negatively) charged center to which an electron (hole) is bound by Coulomb interaction. So we are faced with a problem similar to that of a hydrogen atom, leading, in the simplest approximation, to a series of states with binding energy

$$E_{\mathrm{b}}^{\mathrm{D,A}} = \mathrm{Ry}\frac{m_{\mathrm{e,h}}}{m_0}\frac{1}{\varepsilon^2}\frac{1}{n_{\mathrm{B}}^2}\,, \tag{18.2a}$$

where Ry is the *Rydberg energy* of the H atom (Ry = 13.6 eV), n_{B} the main quantum number and ε a dielectric constant. In Fig. 18.1 we show only the states for $n_{\mathrm{B}} = 1$. Depending on the material parameters, one usually finds in effective-mass approximation values for the shallow donor and acceptor binding energies of $5\,\mathrm{meV} \leq E_{\mathrm{b}}^{\mathrm{D}} \leq 50\,\mathrm{meV}$ and $20\,\mathrm{meV} \leq E_{\mathrm{b}}^{\mathrm{A}} \leq 200\,\mathrm{meV}$.

The binding energies of deep donors and acceptors deviate significantly from these values. For a listing of donor and acceptor binding energies in various semiconductors see e.g., [06S1, 82L1].

The radius of the $n_{\mathrm{B}} = 1$ state is given again in effective-mass approximation by

$$a_{\mathrm{lattice}} < a_{\mathrm{D,A}} = a_{\mathrm{B}}\varepsilon\frac{m_0}{m_{\mathrm{e,h}}} \tag{18.2b}$$

ranging from 1 to 20 nm depending again on the material parameters. The fact that these values are larger than the lattice constant justifies the use of the effective-mass approximation. Details concerning which value of ε has to be used are similar to those for excitons, as discussed in Sect. 20.3. In addition, there is some smaller influence of the chemical nature of the atom forming the donor or acceptor which is known as the *central-cell correction* or *chemical shift*.

Similar to the localized phonon modes, the wave function of a shallow donor (or acceptor) can be described as a superposition of Bloch states

$$\phi(\boldsymbol{r}) = \sum_k a_k \phi_k^{\mathrm{CB}}(\boldsymbol{r}). \tag{18.3}$$

The range of $\boldsymbol{k}$ from which significant contributions can be expected increases with decreasing radius in (18.1b), as shown and discussed already with Figs. 13.3 and 18.1b.

The occupation probability of donors and/or acceptors in thermodynamic equilibrium is again governed by Fermi statistics. However, some correction terms appear

$$f_{\mathrm{D/A}} = \frac{1}{1 + 1/g_i \exp\left\{\left(E_{\mathrm{D/A}} \mp E_{\mathrm{F}}\right)/k_{\mathrm{B}}T\right\}} \tag{18.4}$$

where the g_i describe the degeneracy of the level and take care of the fact, that, e.g., an ionized donor D^+ offers two states for electrons with opposite spin in the ground state. However, if this state is occupied by one electron, the level for the second one is considerably blue shifted due to Coulomb repulsion. For details see, e.g., [87B1].

To summarize this subsection we show schematically in Fig. 18.2 the density of states including some impurity centers, the occupation probability in thermodynamic equilibrium and the resulting density of electrons and holes per unit energy. This sketch is not to scale and attention should be payed to the following comments:

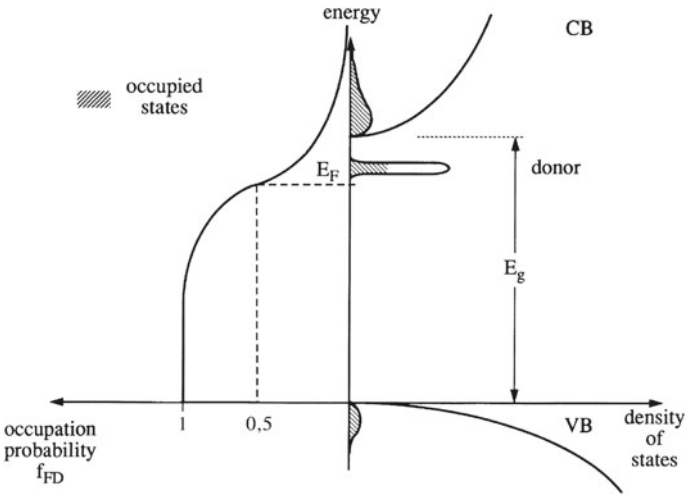


Fig. 18.2 The square-root density of states DOS in the effective-mass approximation of an idealized three-dimensional semiconductor containing some impurities, here a donor state, and the resulting population (*hatched areas*). Energies and occupation are not to scale! See text for further explanation

Usually the relation $k_B T \ll E_g$ holds and the Fermi-function looks more step-like with the consequence that the populations in conduction and valence bands cannot be shown simultaneously on the same linear scale. The density of holes in an n-doped semiconductor is actually very tiny in accordance with global charge neutrality and with the mass-action law (16.26). In the case of shallow donors and room temperature (the typical condition for many electronic devices) the donors are fully depleted yielding a respective density of electrons in the conduction band. Equivalent considerations apply to p-doping.

18.1.2 Shallow Defect Related Radiative Transitions

Pairs of donors and acceptors that are so close in space that their wave-functions overlap are known as *donor–acceptor pairs* (see Fig. 18.1). Actually one can also consider these pairs as simplest version of a *poly-centric bound-exciton complex*. Due to the wave-function overlap the electron on the donor and the hole on the acceptor can recombine. This process is, at least partly, radiative giving rise to a characteristic emission feature.

The energy of the photon resulting from the radiative recombination is given by

$$\hbar\omega_{DA} = E_g - E^b_{D^0} - E^b_{A^0} + \frac{e^2}{4\pi\varepsilon\varepsilon_0 r_{DA}} - m\hbar\omega_{LO} \tag{18.5}$$

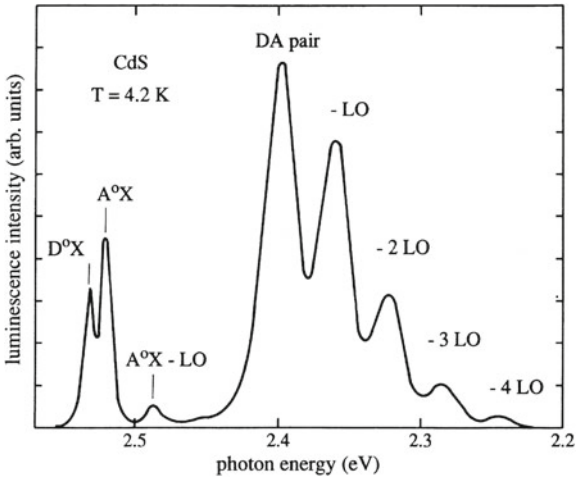


Fig. 18.3 Donor–acceptor pair luminescence in CdS including some phonon replica. The lines around 2.5 eV are related to excitons bound to donors or acceptors [79B1]

for singly charged centers if we neglect the energetic shift caused by the overlap of the wavefunctions of the neutral donor and acceptor. $E^{\mathrm{b}}_{\mathrm{D}^0\mathrm{A}^0}$ are the binding energies of electron and hole to their respective centers.

The forth term on the r.h.s of (18.5), which depends on the distance r_{DA} of the centers, reflects the Coulomb energy of the ionized centers after the recombination. The last term describes *LO phonon replica*. We give examples in Figs. 18.3 and 22.2b.

If the donors and acceptors are introduced substitutionally in a rather simple (e.g., cubic) lattice, then only discrete, well-defined values for r_{DA} are possible, to nearest neighbors and next-nearest neighbors, etc. Consequently using high spectral resolution one sometimes observes that the zero-phonon band consists of discrete lines corresponding to the discrete values of r_{DA} [79C1, 79D1].

With increasing pump intensity, the number of occupied donor and acceptor centers increases and their average distance r_{DA} necessarily decreases. As a consequence one finds that the emission maximum of the pair-band shifts to the blue with increasing excitation due to the Coulomb term in (18.5). This is a very characteristic feature of donor–acceptor pair recombination. It allows distinguishing excitation-induced from thermal effects, which usually result in a red shift. For examples see [80B1, 84V1, 10M1, 10W1].

Another recombination process connected with neutral donors or acceptors is the so-called *free-to-bound transition* [66C1, 03K1, 04A1, 08B1, 09T1]. In this case a free electron or hole recombines with a neutral acceptor or donor, respectively, as shown schematically in Fig. 18.1. The corresponding emission peak is at

$$\hbar\omega_{\mathrm{FB}} = E_{\mathrm{g}} - E^{\mathrm{b}}_{\mathrm{D}^0/\mathrm{A}^0} - m\hbar\omega_{\mathrm{LO}}\,. \tag{18.6}$$

Obviously $\hbar\omega_{\mathrm{FB}}$ is blue-shifted as compared to $\hbar\omega_{\mathrm{DA}}$ and often both processes overlap together with their phonon replica to form one complex or broadened luminescence

spectrum. Examples are shown later in Fig. 22.2c both for donor–acceptor pairs (D^0, A^0) (the index C means a carbon acceptor) and free electron to bound hole transitions (e,A^0).

Since the $\boldsymbol{k}$-conservation is relaxed at the defect, free electrons can recombine with the hole of the acceptor independent of their kinetic energy and vice versa. This fact allows to monitor spectrally the distribution of the free carriers in their band.

In [69M1] the IR luminescence has been observed, which results from the transition of an electron from the conduction band or from an excited state of the donor to the donor ground state. In contrast to the free-to-bound transitions above, which take place across the gap, this process remains limited energetically to one band and its related bound states.

18.1.3 Deep Defects and Related Radiative Transitions

Apart from the shallow donor and acceptors there are deep donors and acceptors as mentioned already above and a variety of other *deep centers*. These are atoms which have one or more energy levels somewhere around the middle of the gap. See e.g., [82L1, 01H1].

For deep centers an approach as in (18.2) is not adequate. The wave function is better described by the parent atomic orbitals, modified by the influence of the surrounding atoms, i.e., by the symmetry of the arrangements of the neighbors. Examples of this type of center are copper, sodium, the rare earths, and the transition metals. They give rise to the green, orange and red emission bands of wide-gap semiconductors such as CdS, ZnO and ZnS [67S1, 76T1, 01L1, 10H1, 10M1].

Carriers in deep centers can sometimes couple strongly to bulk and/or localized phonon modes giving rise to very broad emission (and absorption) features with Huang-Rhys factors S significantly larger than 1. We give an example in Fig. 18.3. The phonon structure, which is characteristic for this luminescence band allows determining the Huang-Rhys factor for this Cu centre in ZnO to be 6.5.

The assignment of the emission in a certain spectral range is frequently not unique. The well known green luminescence of ZnO can, e.g., be caused by oxygen vacancies or Cu-centers [01L1, 10K2]. In the latter case, which is shown in Fig. 18.4, it is accompanied by clear phonon structures.

Some deep centers can exchange carriers with both the conduction and the valence band (in contrast to the deep and shallow donors and acceptors). In this case they are called *recombination centers*. The recombination can be radiative or non-radiative. If this recombination is fast and non-radiative, these centers are known as “luminescence killers”. They are to a large extent responsible for the low luminescence yield of many semiconductors. Iron is an example of such a center.

A group of (deep) centers is often formed by so-called *isoelectronic traps*. These are frequently atoms of the same column of the periodic table as the one which they substitute, i.e., they have the same electron configuration in the outer shell. An example of an isoelectronic trap would be Te replacing S or Se in ZnS or ZnSe.

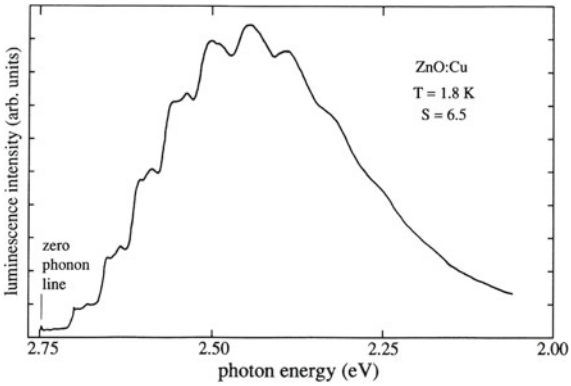


Fig. 18.4 Deep center emission of ZnO:Cu (reprinted with permission from [81K1] ©1981 Elsevier)

Some centers have various levels in the forbidden gap and so transitions within the center can be investigated as in Cu, Ni or other ions. Since the chance of having various levels in the gap increases with increasing width of the gap, such *internal transitions* are best investigated in wide-gap semiconductors and in insulators. The later are efficient solid state laser materials, e.g., when doped with Cr, Ti or Nd. For more details on defects in insulators we refer the reader to [81A1] and references therein.

18.1.4 Defects and Doping in Quantum Wells

In connection with (M)QW two special methods of doping should be mentioned: *δ-doping* means the introduction of a two-dimensional sheet of doping atoms during epitaxial growth. The concentration in the growth direction then has an almost δ-function-like profile. *Modulation doping* means introducing the doping atoms into the barriers of a (M)QW. The electrons (holes) are thermally ionized into the conduction (valence) band of the barrier, reach the well by diffusion, and are captured in it. This allows the production of high two-dimensional carrier densities in the well with high mobility, since the charged impurities are separated spatially from the mobile carriers. Partly even an additional undoped barrier layer is added between the doped barrier and the quantum well, thus reducing considerably the scattering.

The space charges of the ionized doping atoms and of the free carriers lead to a characteristic curvature of the bands which can be calculated by solving the Poisson equation

$$-\Delta\phi = \frac{\rho}{\varepsilon\varepsilon_0}, \tag{18.7}$$

where ϕ is the electrostatic potential and ρ the space charge as shown schematically in Fig. 18.5. Modulation doping is the basis of devices called high-electron- mobility transistors (HEMT) or modulation-doped field-effect transistors (MODFET).

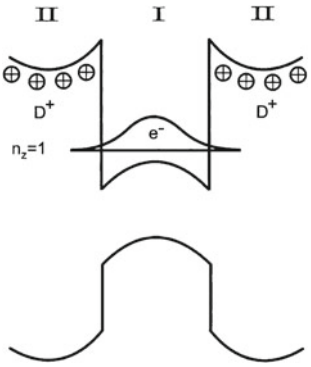


Fig. 18.5 The curvature of the band edges in the case of symmetric modulation doping of a MQW (schematic) including the first quantized electron level

The binding energy of donors and acceptors in SL and (M)QW depends in addition to the parameters (18.2) on the distance from the barrier [86G1]. Furthermore, it has been found in quantum structures that the doping efficiency depends on the crystallographic surface on which the structure has been grown [05E1].

More information on defects in bulk materials and in MQW structures is given e.g., in [88P1, 91E1] and references therein.

18.2 Disordered Systems and Localization

In this section we want to briefly summarize some consequences of structural disorder in semiconductors on their electronic and optical properties. Disorder is actually a widely present phenomenon in solid-state physics which has been studied extensively. More exhaustive treatments of this topic, from which the following discussion is largely extracted, are [79M1, 79Z1, 81A1, 83Z1, 84S1, 85K1, 85S1, 86L1, 00A1]. Furthermore we want to mention the pioneering work of P. Anderson [58A1].

18.2.1 Anderson Model of Localization

We start by recollecting some features of a periodic potential as is shown schematically in Fig. 18.6a. This potential is known as the *Kronig–Penney potential*. Due to the tunneling of the wave-function from the wells into the barriers, there are finite overlap integrals and consequently a band of Bloch states with a certain width B arises. We have considered similar scenarios with (15.7) or with Figs. 17.6 and 17.8.

We now want to introduce disorder into this system. This can be done by varying the depths of the potentials statistically within a width V_0 (*diagonal disorder*) or by varying the widths of the potential wells and barriers and thus the coupling (*off-diagonal disorder*). The first case leads to the so-called *Anderson model*, the second

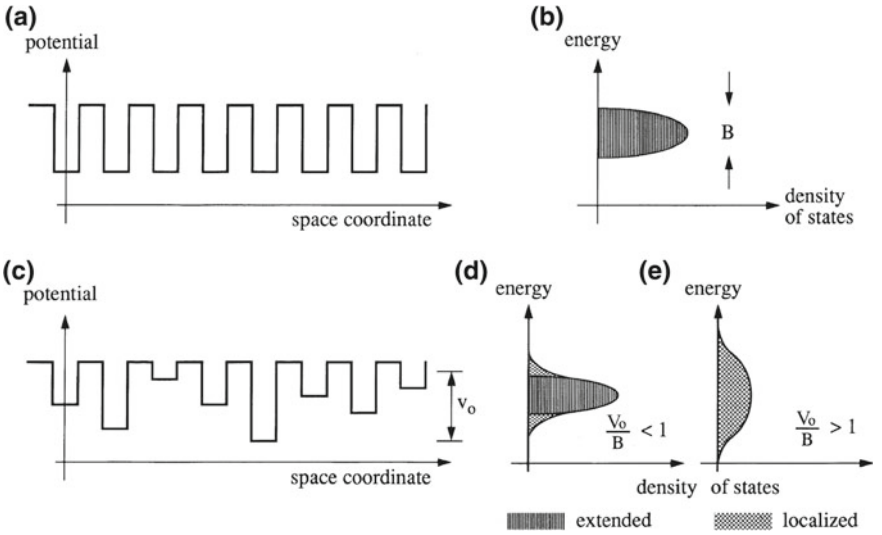


Fig. 18.6 A periodic potential (**a**) and the resulting band, represented by its density of states (**b**), a disordered system (**c**) and bands with localized and delocalized states for small fluctuation width V_0 (**d**) and with localized states only (**e**)

to the *Lifshitz model*. In practice, both aspects of disorder will occur simultaneously, but from the theoretical point of view it is already difficult to treat them individually.

We want to briefly outline some ideas of the Anderson model which are illustrated in Figs. 18.6 and 18.7. If some diagonal disorder is introduced, as in Fig. 18.6c, two things happen—the sharp edges of the density of states are smeared out typically by exponential tails, and a new type of eigenstate appears, namely localized states. If

$$B V_0^{-1} > 1 \,, \tag{18.8a}$$

then there are both, localized states at the band tails and extended states in the center, which, however, are different from Bloch states, as we shall see shortly (Figs. 18.6d and 18.7b, c). For

$$B V_0^{-1} < 1 \,, \tag{18.8b}$$

there are only localized states in the band (Figs. 18.6c and 18.7c).

To explain the concept of extended and localized states we show in Fig. 18.7a a Bloch wave. This is a superposition of atomic orbitals ϕ_n placed at every lattice site $\boldsymbol{R}_i$ with a proper long-range phase correlation, corresponding to the LCAO method mentioned in Sect. 15.1.2:

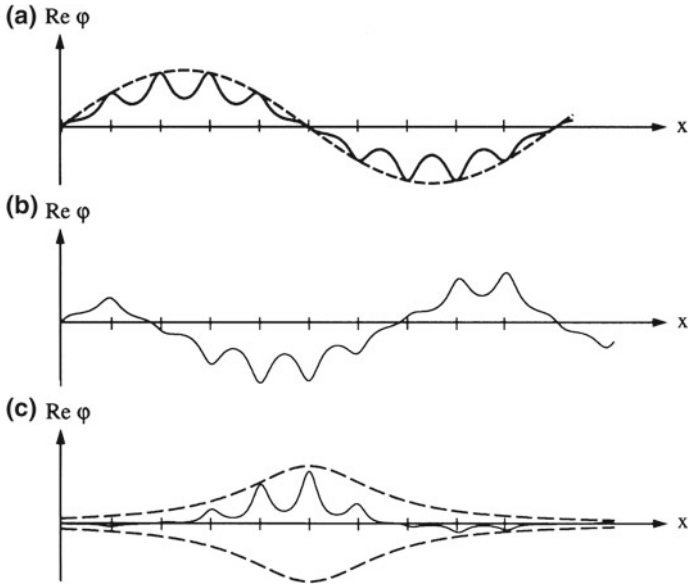


Fig. 18.7 Schematic drawing of the real part of the wave function for a periodic potential (**a**) and for a disordered system of an extended state (**b**) and of a localized one (**c**)

$$\begin{aligned}\phi_{\boldsymbol{k}}^{\text{Bloch}}(\boldsymbol{r}) &= \sum_i \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_i}\phi_n(\boldsymbol{r}-\boldsymbol{R}_i)\\ &= \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\sum_i \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{R}_i)}\phi_n(\boldsymbol{r}-\boldsymbol{R}_i)\\ &= \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}u_k(\boldsymbol{r})\,. \end{aligned} \tag{18.9a}$$

In an extended state in a disordered system (Fig. 18.7b) the long-range phase correlation is lost resulting in

$$\phi^{\text{ext}}(\boldsymbol{r}) = \sum_i c_i\phi_n(\boldsymbol{r}-\boldsymbol{R}_i)\,. \tag{18.9b}$$

A localized state obtains an envelope which is centered around a localization site $\boldsymbol{r}_0$ and which decays exponentially with a *localization length* ξ (Fig. 18.7c)

$$\phi^{\text{loc}}(\boldsymbol{r}) = \sum_i d_i\phi_n(\boldsymbol{r}-\boldsymbol{R}_i)\mathrm{e}^{-|\boldsymbol{r}-\boldsymbol{r}_0|/\xi}\,. \tag{18.9c}$$

The energy which separates extended from localized states is called *mobility edge*. If the carriers were to have an infinite phase-relaxation time T_2 (see the second volume of this book for the explanation of this quantity) in their states and thus also an infinite lifetime, it is easy to show that the mobility edge is sharp: if extended and localized

states were to coexist in a certain energy range, then the slightest perturbation would mix these states and consequently every localized state would obtain some admixture of an extended wave function (and vice versa) transforming it to an extended state. Since the T_2 times of the carriers in their states are limited, the mobility edge is usually a transition region of a certain width.

Another important consequence of disorder is the following. A disordered system is no longer invariant with respect to translations. As a consequence, the wavevector is no longer a good quantum number and the dispersion relation $E(\boldsymbol{k})$ loses, strictly speaking, its meaning. A disordered system is in principle characterized only by its density of states $D(E)$.

As usual, there are exceptions to this rule: excitations with a wavelength λ long compared to the typical length scale of the disorder fluctuations "average" over the disorder and can be characterized by a wave vector $\boldsymbol{k}$ in the sense of a continuum approximation. This is fulfilled for example for long wavelength (acoustic) phonons or photons. The latter case is easily checked by looking through a glass window or into clear water. See also Sect. 8.3.2.

Weak disorder, e.g., as defined by (18.8a), will produce some localized states at the band edges ($10^{16} - 10^{18}\,\mathrm{cm}^{-3}$) but the extended states in the band ($10^{22} - 10^{23}\,\mathrm{cm}^{-3}$) will be close to Bloch-type waves. See Fig. 18.7.

A crucial property for localization effects is the dimensionality of the system. It has been shown by general arguments that the slightest disorder will in principle localize all states for dimensions $d \leq 2$. However the localization length [ξ in (18.9c)] can be extremely long and in many cases exceeds the length of the sample used. As a consequence these considerations are largely of theoretical interest and only to a limited extent of practical relevance.

18.2.2 Realizations of Disorder in Semiconductors

After characterizing briefly some properties of disordered systems we now proceed to the inspection of realizations of disorder in semiconductors. We discuss, in roughly increasing magnitude of disorder, heavily doped semiconductors, alloy semiconductors, quantum-well structures with well-width fluctuations, and amorphous semiconductors.

In Fig. 18.8a, we show the density of states for the conduction band of a semiconductor containing donors of variable concentration. If the concentration of donors is so low that the wave-functions do not overlap, i.e., $a_\mathrm{B}^3 N_\mathrm{D} \ll 1$, where N_D is the concentration of donors and a_B the radius of the wave function according to (18.2b), then their density of states is δ-function like. The state with $n_\mathrm{B} = 1$ for a donor with one electron only is show in Fig. 18.8a. A donor may bind a second electron with opposite spin. As a result of the Coulomb interaction the energy for the two electrons will be higher. This is indicated by the second peak. The states with $n_\mathrm{B} > 1$, which the electron could also occupy, are again neglected.

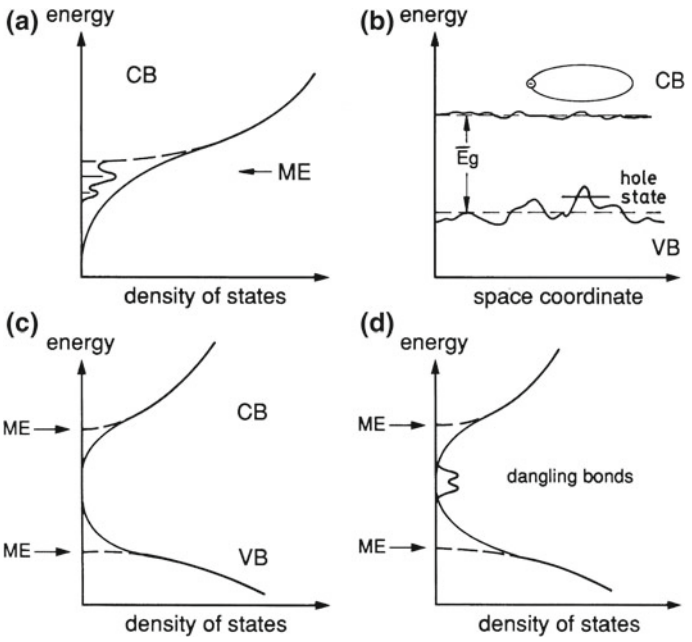


Fig. 18.8 The appearance of tails of localized states due to increasing doping (**a**), potential fluctuations, e.g., in an alloy (**b**, **c**) or in amorphous semiconductors (**d**)

If N_D is increased, the wave functions of the start to overlap and this again results in the formation of a so-called *impurity band*. Its width increases with N_D and eventually merges with the conduction band, forming an (exponential) tail of the density of states. Depending on the doping and possibly on compensation through acceptors, the Fermi level can be situated for $T \to 0$ above or below the mobility edge. In the first case, the system is called a metal, since it has finite conductivity at $T \to 0$, in the second it is an insulator. The transition from an insulator to a metal with increasing doping is also known as *Mott transition*. There was a long discussion as to whether this transition is continuous or not, i.e., whether there is something like minimum metallic conductivity in the second case. Experiments, e.g., with phosphorous-doped silicon (Si:P), indicate that there is no minimum conductivity [84S1, 85K1, 86L1].

The next case of disorder is intrinsic to alloy semiconductors like $Ga_{1-y}Al_yAs$, $CdS_{1-x}Se_x$, $ZnSe_{1-x}Te_x$ or $Zn_{1-y}Cd_yS$. Many, but not all of them can be grown for every value of x or y between 0 and 1, others like $Ga_{1-y}In_yN$ or $GaAs_{1-x}N_x$ have a miscibility gap. They usually have a well-defined crystal structure (e.g., of the zincblende or of the wurtzite type). The disorder is introduced by the statistically distributed occupation of anion (or cation) sites with the two different atoms.

Even though an alloy has no translational invariance one can create it artificially in the so-called *virtual crystal approach*. One assumes that the unit cell contains atoms with a mixture of the properties according to x or y. This approach explains

for example the often observed continuous shift of E_g with composition (Fig. D.2). This shift can be described by [99K1, 00P1, 04K1]

$$E_g(x) = E_g(x=0) + \big(E_g(x=1) - E_g(x=0)\big)\big((1-b)x + bx^2\big) \tag{18.10}$$

where the usually small bending or *bowing parameter* b describes the deviation from a linear relation. In some cases, like in $In_{1-y}Ga_yAs_{1-x}N_x$ the bowing parameter is so large, that the system goes with increasing x through a band-gap minimum which is situated considerably below that of $x=0$ or 1 [07H1].

The fluctuations now arise from the fact that the composition varies on a microscopic scale (e.g., the exciton radius, see below). In $CdS_{1-x}Se_x$, for example, there is in some places a little more S, in others more Se. This *compositional disorder* frequently results in the more ionic systems in preferential fluctuation of the valence band for anion substituted materials like $CdS_{1-x}Se_x$, or of the conduction band for cation-substituted materials like $Zn_{1-y}Cd_yS$. In the more covalent materials like $Al_{1-y}Ga_yAs$ both bands fluctuate with composition. These fluctuations can again lead to localized states. The *tailing parameter* E_0, which describes the exponential tail in

$$D(E) = \frac{N_0}{E_0} e^{-E/E_0}, \tag{18.11}$$

where N_0 is the total number of tail states, depends on the material parameters and on the composition. It is obvious that E_0 disappears for x, $y=1$ and x, $y=0$, that it increases for increasing $E_g(x=1) - E_g(x=0)$ and for decreasing effective mass of the particle which is localized. The latter point depends on the fact that in three dimensions a potential well must have a certain width R and depth V to localize or bind a particle of mass in m with

$$V R^2 \geq \frac{\hbar^2}{m_{e,h}} \tag{18.12}$$

and explains why holes are more easily localized than electrons.

For the dependence of E_0 on x one finds different but similar formulas in literature. We give here one according to [78B1]. For discussions of this topic concerning also excitons see [78B1, 84S1, 91K1, 92P1, 93K1, 99K1, 00P1].

$$E_0 = \frac{1}{178}\beta^4 x^2 (1-x)^2 m_{e,h}^3 \left(\hbar^6 N^2\right)^{-1}, \tag{18.13}$$

where N is the density of atoms and β the derivative of the position of the edge of the band with respect to x. Often one finds from experiment a slightly asymmetric behavior of $E_0(x)$ caused by the term β and other effects.

Electron–hole pairs that have been excited non-resonantly in extended band states relax towards the localized states and can get trapped as whole excitons (see

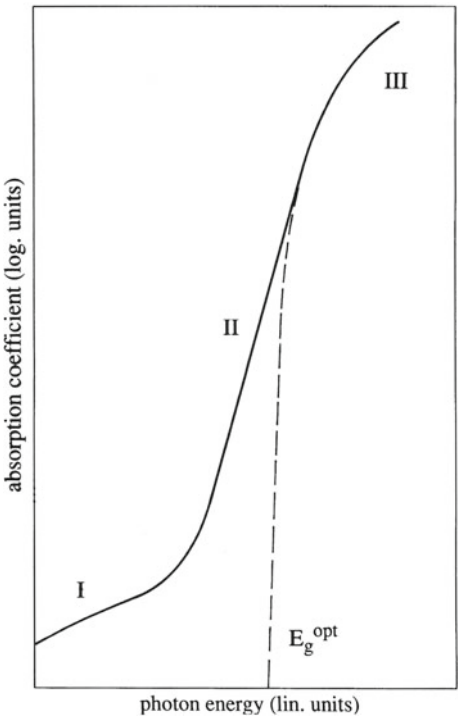


Fig. 18.9 The various regimes of the absorption coefficient in disordered semiconductors (repreinted with permission from [72W1] ©1972 American Physical Society https://doi.org/10.1103/PhysRevB.5.3144)

Sect.22.2) in semiconductors with large exciton binding energy. In strongly disordered semiconductors and small exciton binding electrons and holes get localized individually. The typical shape of the absorption spectrum of such a strongly disordered (e.g., amorphous) system is shown in Fig. 18.9.

In region I the absorption is weak and is caused by impurities in the alloy or amorphous semiconductor, i.e., atoms of a different chemical nature.

Then region II follows where the absorption coefficient increases exponentially with photon energy. This regime is also known as the *Urbach tail*. Its origin in disordered semiconductors is the exponential tail of the density of states of localized excitons or carriers which can be reached without $\boldsymbol{k}$-selection.

Above follows region III comprising the absorption via transitions into extended states. An extrapolation of this part to lower energies, e.g., with a $(E'-E_0)^2$ law, allows one to define an *optical* or *mobility gap*. This region is called called *Tauc regime* [89S1, 89T1]. The optical gap can also be found by excitation-spectroscopy (see 23.2) of the photo-current.

Absorption spectra corresponding to the schematic drawing of Fig. 18.9 have been found in amorphous Si, in chalcogenide glasses, and in alloy semiconductors such as $CdS_{1-x}Se_x$ and $ZnSe_{1-x}Te_x$. Examples are presented, e.g., in [72W1, 89T1, 99R1, 04K1]. Please note that a fit according to the Tauc regime or according to a simple band-to-band transition is not suitable for semiconductors with strong electron–hole

Coulomb correlation (leading to a large exciton binding energy) like ZnO or its alloys like $Mg_{1-x}Zn_xO$. For a detailed discussion see [10K1, 10K2].

The here descried gradual increase of the absorption, and via Kramers–Kronig transformation also of the real part of the reflective index, in many cases even prevents the observation of discrete exciton resonances in reflection.

The analysis of the Urbach part of the absorption spectrum gives the *tailing parameter* ε_{loc} describing the density of localized states below the mobility edge E_{ME}.

$$N(E) = \frac{N_0}{\varepsilon_{\text{loc}}} \mathrm{e}^{-\mathrm{E}/\varepsilon_{\text{loc}}} \quad E < E_{\text{ME}} \tag{18.14}$$

In Fig. 18.10 we show ε_{loc} for $CdS_{1-x}Se_x$ as a function of x. This quantity necessarily vanishes for the ordered binary compounds $x = 0$ and $x = 1$, see also Sect. 22.2. In between it goes through a maximum with a shape that is not symmetric with respect to $x = 0.5$, partly since the quantity $\mathrm{d}E_{\mathrm{g}}/\mathrm{d}x$ changes with x, which enters in (18.10), and (18.13).

We now consider disorder in systems of reduced dimensionality like MQW. Actually there are two origins of disorder. One is the alloy broadening or disorder felt by the carriers if the well or the barrier or both are made from an alloy or if an alloy forms at the boundary due to inter-diffusion or segregation. Another is due to well-width fluctuations. Though great progress has been made in growing atomically flat interfaces by modern epitaxial techniques, (see Appendix D) fluctuations of l_z of at least the order of one atomic (or molecular) layer are hardly avoidable. Due to the

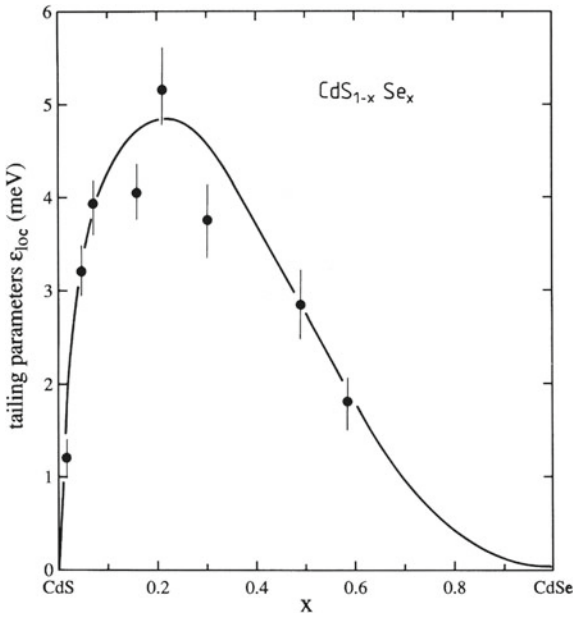


Fig. 18.10 The dependence of the tailing parameter on alloy composition (reprinted with permission from [92P1] ©1992 Elsevier)

dependence of the quantization energy on l_z, carriers (and excitons) can be localized or trapped at regions with larger l_z.

The continuous transition from these localization sites to self-assembled quantum islands has already been mentioned in Sects. 17.5 and Appendix D. For examples of localization in two dimensions see, e.g., [85K1, 86L1, 00P1] and references therein.

The strongest disorder occurs in *amorphous semiconductors*, the prominent example being amorphous or α-Si. In these systems only short-range order remains, e.g., a tetrahedral coordination of most but not of all Si atoms. But there is no long-range order at all caused by the facts that the coordination of some atoms is different from the one in a perfect crystal and that the bond lengths and the angles between the bonds show some fluctuations. If we start in a $CdS_{1-x}Se_x$ crystal at a S atom and proceed by a lattice vector $\boldsymbol{R}$ we will hit an atom. We are not sure, however, whether it will be S or Se. In an α-Si there is only an average bond length but nothing like a lattice "constant" and if we move from a Si atom by a vector $\boldsymbol{R}$ of crystalline Si we are not even sure whether we will find an atom at all.

To illustrate this concept, we show in Fig. 18.11 a hypothetic two-dimensional crystal with hexagonal structure and threefold coordination and in (b) the corresponding amorphous material in which some five and seven atom rings occur and also some dangling bonds.

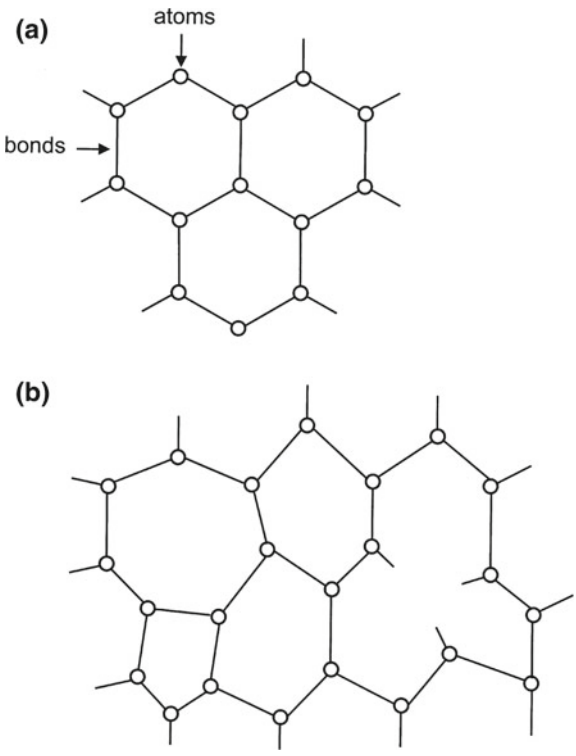


Fig. 18.11 A hypothetical two-dimensional crystalline (**a**) and amorphous (**b**) material

This strong disorder leads to substantial exponential tails of the densities of states both for conduction and valence bands, which cover the whole "forbidden gap" (Fig. 18.8d). Unsaturated covalent bonds in Si, so-called *dangling bonds*, form states situated energetically close to the center of the gap since it is undetermined whether they would form a bonding or an anti-bonding orbital with the next Si. The states of the dangling bonds can be saturated by hydrogen doping, opening interesting technological applications for α-Si:H. A short selection of books and conference proceedings for the properties of amorphous semiconductors is [74A1, 84H1, 84T1, 88W1].

18.2.3 Weak Localization and Percolation

To conclude this section, we now outline the concepts of *weak localization* by enhanced back-scattering and of percolation. We come back to these approaches in Sect. 22.2.

In the first case we consider an otherwise perfect crystal which contains randomly distributed elastic scattering centers (see Fig. 18.12). To obtain the probability $w_{\mathrm{A}\to\mathrm{B}}$ that a particle, e.g., a carrier starting from A reaches B we have to sum over the amplitudes $a_i^{\mathrm{A}\to\mathrm{B}}$ of all possible ways to come from A to B and to square the sum:

$$w_{\mathrm{A}\to\mathrm{B}} \sim \left|\sum_i a_i^{\mathrm{A}\to\mathrm{B}}\right|^2 = \sum \left|a_i^{\mathrm{A}\to\mathrm{B}}\right|^2 + \sum_{i,j} a_i^{\mathrm{A}\to\mathrm{B}} a_j^{\mathrm{A}\to\mathrm{B}} \approx \sum \left|a_i^{\mathrm{A}\to\mathrm{B}}\right|^2 . \tag{18.15a}$$

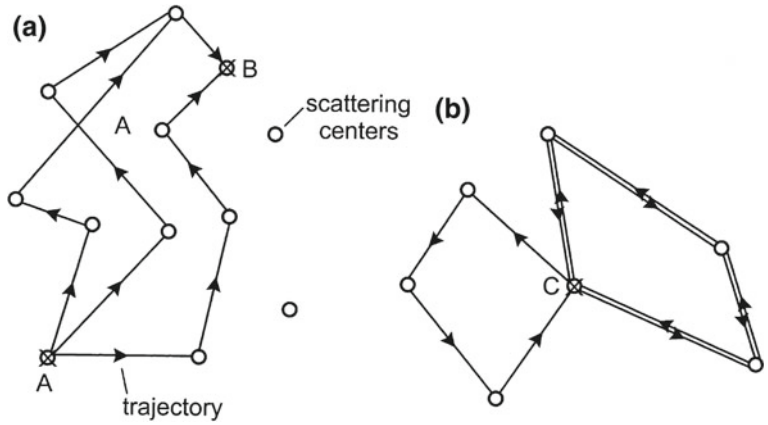


Fig. 18.12 Various possibilities for a particle to come from A to B or to be back-scattered from C to C. The *open dots* symbolize scattering centers

The mixed terms all have different relative phases and cancel essentially as a sum. In the case of back-scattering we have to start in the same way, i.e.,

$$w_{\mathrm{C}\to\mathrm{C}} = \left|\sum_i a_i^{\mathrm{C}\to\mathrm{C}}\right|^2 = \sum_i \left|a_i^{\mathrm{C}\to\mathrm{C}}\right|^2 + \sum_{i,j} a_i^{\mathrm{C}\to\mathrm{C}} a_j^{\mathrm{C}\to\mathrm{C}} . \tag{18.15b}$$

The difference compared to (18.15a) is that there not only remains the term $\left|\sum_i a_i^{\mathrm{C}\to\mathrm{C}}\right|^2$ but also the mixed terms. The latter are made up from $a_i^{\mathrm{C}\to\mathrm{C}}$ and their time-reversed parts shown for one case in Fig. 18.12, since the total phase shift encountered in both directions is equal when only *elastic* scattering is assumed. As a consequence we obtain

$$w_{\mathrm{C}\to\mathrm{C}} = 2\left|\sum_i a_i^{\mathrm{C}\to\mathrm{C}}\right|^2 . \tag{18.15c}$$

This result means that back-scattering is by a factor of 2 more probable than scattering from $\mathrm{A}\to\mathrm{B}\neq\mathrm{A}$.

It should be noted that similar arguments can also be formulated in $\boldsymbol{k}$-space. Due to elastic scattering only the length of $\boldsymbol{k}$ remains constant and the scattering from an initial $\boldsymbol{k}_i$ to $\boldsymbol{k}_f = -\boldsymbol{k}_i$ is again by a factor of 2 more probable than scattering to any other $\boldsymbol{k}_f \neq \boldsymbol{k}_i$.

This *enhanced back-scattering* is obviously a beginning of localization. The concept has been verified for scattering of photons from small polystyrole spheres [85A1, 85W1, 86A1, 86S1] and in electric conductivity. In the latter case there are also considerable contributions from inelastic carrier scattering to the resistivity. The part due to enhanced back-scattering can be identified by applying a magnetic field $\boldsymbol{B}$, since the simple time reversal argument is no longer valid for $\boldsymbol{B}\neq 0$. For more details see, e.g., [84B1]. We will come back to the localization of light in connection with random lasing (see second volume).

The concept of *percolation* will be explained with a simple-cubic lattice in Fig. 18.13. Some of the unit cells are randomly painted black. All neighboring cells that are painted black are said to belong to one cluster. We see in Fig. 18.13 clusters consisting of one, two, three and seven unit cells. With increasing occupation probability (i.e., increasing number of black unit cells) the average cluster size growths. Eventually a cluster appears that extends from $-\infty\to+\infty$. This is the *percolating cluster* that corresponds to an extended state, while finite clusters are "localized". The occupation probability at which the first percolating cluster appears is the critical occupation f_c.

It is easy to understand that f_c depends on the dimensionality (in a one-dimensional case f_c is trivially equal to unity), on the type of lattice e.g., cubic, hexagonal, or triangular and on the fact if we consider site or bond percolation. For details see, e.g., [85S1].

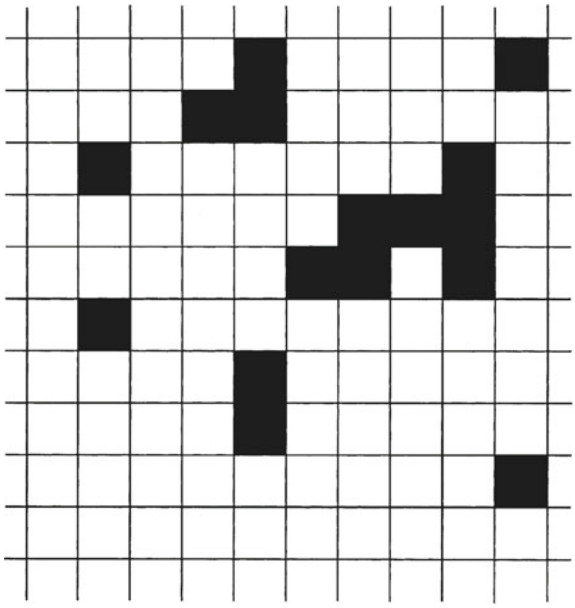


Fig. 18.13 Cluster formation in a simple cubic lattice in two dimensions

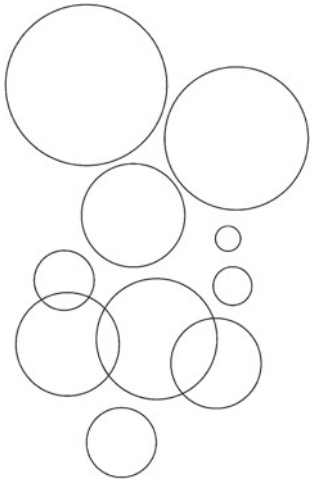


Fig. 18.14 Clusters of overlapping spheres

One possibility to make the percolation model more realistic is e.g., the following one, which has been used with some more sophistication to successfully explain the optical properties of excitons in alloys like $CdS_{1-x}Se_x$ (see [99K1] and Sect. 22.2). We assume that a localized state like the hole state in Fig. 18.8b is caused by a cluster of Se atoms in $CdS_{1-x}Se_x$ and that the wave function localized at this Se cluster has a spherical shape as indicated in a two-dimensional analogue (which corresponds to quantum wells [00P1]) in Fig. 18.14.

Wave functions that overlap spatially form a cluster. At a certain density of spheres a percolating cluster will appear that corresponds to an extended state while finite clusters again correspond to localized ones. The density and the size of the spheres can be influenced, e.g., by varying the energy in the tail of the density of states. The appearance of a percolating cluster with increasing energy corresponds then to the mobility edge.

An demonstration experiment showing localization in a disordered chain of macroscopic, coupled pendula has been described and verified experimentally in [89K1].

We will end the discussion of the properties of electronic states in semiconductors here and concentrate in the next chapters on electron–hole pairs and excitons, which are actually the relevant complexes for optical properties, but which cannot be understood without knowledge of the properties of the constituting carriers.

Problems

18.1 Calculate the binding energy of electrons (holes) to donors (acceptors) for some of the common semiconductor materials. Compare with experimental data as given, e.g., in [82L1]. Deduce the order of magnitude of the central cell correction. Check the ratio of the binding energy and of the energy of the LO phonon in the respective material, and try to anticipate the consequence for the choice of ε.

18.2 Calculate the Bohr radii of acceptors and donors for the extreme values found in the first problem. How many unit cells or atoms are contained in the volume of a donor?

18.3 Calculate the shift in energy for the donor–acceptor pair recombination when donor and acceptor are nearest possible neighbors in a zinc-blende lattice, and when they have a separation of three, and of ten lattice constants. Can one observe emission from pairs with a separation $d \gg a_B \varepsilon \frac{m_0}{m_c}$.

18.4 What is a dislocation? Which types of dislocations do you know or can you find in textbooks?

References

[58A1] P.W. Anderson, Phys. Rev. **109**, 1492 (1958)
[66C1] K. Colbow, Phys. Rev. **141**, 742 (1966)
[67S1] Ch. Solbrig, E. Mollwo, Solid State Commun. **5**, 625 (1967)
[69M1] I. Melngailis et al., Phys. Rev. Lett. **23**, 1111 (1969)
[72W1] D.L. Wood, J. Tauc, Phys. Rev. B **5**, 3144 (1972)
[74A1] J. Stuke, W. Brenig (eds.), *Amorphous and Liquid Semiconductors* (Taylor and Francis, London, 1974 and 1990)

[76T1] E. Tomzig, R. Helbig, J. Lumin. **14**, 403 (1976)
[78B1] S.D. Baranovskii, A.L. Efros, Sov. Phys. Semicond. **12**, 1328 (1978)
[79B1] G. Blattner, Diplom thesis, University of Karlsruhe (1979)
[79C1] K. Cho (ed.), Excitons, *Topics in Current Physics*, vol. 14 (Springer, Berlin, 1979)
[79D1] P. Dean, Excitons, *Topics Current Physics*, vol. 14 (Springer, Berlin, 1979), p. 11
[79M1] N.F. Mott, E.A. Davies, *Electronic Processes in Non-Crystalline Materials*, 2nd edn. (Clarendon, Oxford, 1979)
[79Z1] J.M. Ziman, *Models of Disorder* (Cambridge University Press, Cambridge, 1979)
[80B1] G. Blattner, R. Helbig, C. Klingshirn, Solid State Commun. **33**, 341 (1980)
[81A1] B. Di Bartolo (ed.), *Proceedings of the International School on "Atomic and Molecular Spectroscopy": Energy Transfer Processes in Condensed Matter* (1983), NATO ASI Series B, vol. 114 (Plenum Press, New York, 1984), *Spectroscopy of Solid-State Laser Type Materials* (1985), Ettore Majorana International Science Series, vol. 30 (Plenum Press, New York, 1987); *Optical Properties of Excited States in Solids* (1991), NATO ASI Series B, vol. 301 (Plenum Press, New York, 1994)
[81K1] R. Kuhnert, R. Helbig, J. Lumin. **26**, 203 (1981)
[82L1] O. Madelung, U. Rössler (ed.), *Landolt–Börnstein*. New Series, Group III, vol. 17 a to i, 22 a and b, 41 A to D (Springer, Berlin, 1982–2001)
[83Z1] R. Zallen, *Physics of Amorphous Solids* (Wiley, New York, 1983)
[84B1] G. Bergmann, Phys. Rep. **107**, 1 (1984)
[84H1] W. Heywang, *Amorphe and Polykristalline Halbleiter* (Springer, Berlin, 1984)
[84S1] B.I. Shklovskii, A.L. Efros, *Electronic Properties of Doped Semiconductors*, Springer Series in Solid-State Sciences, vol. 45 (Springer, Berlin, 1984)
[84T1] J.D. Jonnopoulos (ed.), *The Physics of Hydrogenated Amorphous Silicon* (Springer, Berlin, 1984)
[84V1] I.K. Vereshchagin, V.A. Nikitenko, S.G. Stoyukhin, J. Lumin. **29**, 215 (1984)
[85A1] M.P. Albada, A. Lagendijk, Phys. Rev. Lett. **55**, 2692 (1985)
[85K1] B. Kramer, G. Bergmann, Y. Bruynserade (eds.), *Localization, Interaction and Transport Phenomena*, Springer Series in Solid-State Sciences, vol. 61 (Springer, Berlin, 1985)
[85S1] D. Stauffer, *Introduction to Percolation Theory* (Taylor and Francis, London, 1985)
[85W1] P.E. Wolf, G. Maret, Phys. Rev. Lett. **55**, 26969 (1985)
[86A1] E. Akkermans, P.W. Wolf, R. Maynard, Phys. Rev. Lett. **56**, 1471 (1986)
[86G1] R.L. Green, K. Bajaj, Phys. Rev. B **34**, 951 (1986)
[86L1] D.M. Finlayson (ed.), *Localization and Interaction* (SUSSP, Edinburgh, 1986)
[86S1] M. Stephen, Phys. Rev. Lett. **56**, 1809 (1986)
[87B1] J.S. Blakemore, *Semiconductor Statistics* (Dover Publication, Dover, 1987)
[88P1] C.Y. Fong, I.P. Batra, S. Ciraci (eds.), *Properties of Impurity States in Superlattice Semiconductors*, NATO ASI Series B, vol. 183 (Plenum, New York, 1988)
[88W1] W. Walukiewizc, Phys. Rev. B **37**, 4760 (1988)
[89K1] C. Klingshirn, *Spectroscopy of Solid-State Laser Type Materials* (1985), Ettrore Majorana International Science Series, vol. 30 (Plenum Press, New York, 1987), p. 111
[89S1] E.A. Shiff in *Disordered Solids: Structures and Processes* (1987), Ettore Majorana International Science Series, vol. 46 (Plenum Press, New York, 1989), p. 153
[89T1] P.C. Taylor in *Laser Spectroscopy of Solids II*. Topics in Applied Physics, vol. 65 (Springer, Berlin, 1989), p. 257
[91E1] A.I. Ekimov, Phys. Scr. T **39**, 217 (1991)
[91K1] A.A. Klochikhin, S.G. Ogloblin, Sov. Phys. JETP **73**, 1122 (1991)
[92P1] S. Permogorov, A. Reznitsky, J. Lumin. **52**, 201 (1992)
[93K1] A.A. Klochikhin et al., Phys. Rev. B **48**, 3100 (1993)
[95F1] W. Faschinger, Mater. Sci. Forum **182–184**, 29 (1995)
[99K1] A. Klochikhin et al., Phys. Rev. B **59**, 12947 (1999)
[99R1] A. Reznitsky et al., Phys. Rev. B **59**, 10268 (1999)
[00A1] D. ben Avraham, S. Havlin, *Diffusion and Reactions in Fractals and Disordered Systems* (Cambridge University Press, Cambridge, 2000)

[00K1] E. Kurtz et al., Thin Solid Films **367**, 68 (2000), ibid. **412**, 89 (2002), J. Crystal Growth **214/215**, 712 (2000), Phys. Status Solidi (b) **229**, 519 (2001)
[00N1] S. Nakamura, S. Pearton, G. Fasol, *The Blue Laser Diode: The Complete Story* (Springer, Heidelberg, 2000)
[00P1] S. Permogorov et al., J. Crystal Growth **215/215**, 1158 (2000)
[01H1] Q. Huang et al., J. Crystal Growth **227/228**, 117 (2001)
[01L1] F. Leiter et al., Phys. B **308–310**, 908 (2001)
[02H1] D.M. Hofmann et al., Phys. Rev. Lett. **88**, 045504 (2002)
[03K1] R.Y. Korotkov, M.A. Reshchikov, B.W. Wessels, Phys. B **325**, 1 (2003)
[04A1] B. Arnaudov et al., Phys. Rev. B **69**, 115216 (2004)
[04K1] A. Klochikhin et al., Phys. Rev. B **69**, 085308 (2004)
[05E1] S.C. Erwin, Nature **436**(July issue), 91 (2005)
[06S1] S.M. Sze, K.K. Ng *Physics of Semiconductor Devices*, 3nd edn. (Wiley-Interscience, New York, 2006)
[07H1] M. Hetterich, Habilitation Thesis, Karlsruhe (2007)
[07K1] C. Klingshirn, Phys. Status Solidi B **244**, 3027 (2007)
[08B1] R. Buczko et al., Acta Phys. Polonica, A **114**, 1079 (2008), A IP Conf. Proc. **1199**, 205 (2010)
[08L1] K.-I. Lee et al., Phys. Status Solidi C **5**, 3344 (2008)
[09L1] J.L. Lyons, A. Janotti, C.G. Van der Walle, Appl. Phys. Lett. **95**, 252105 (2009)
[09M1] M.D. McCluskey, S.J. Jokela, J. Appl. Phys. **106**, 071101 (2009)
[09T1] K. Thonke et al., Microelectron. J. **40**, 210 (2009)
[10H1] A. Hoffmann in *Zinc Oxide: From Fundamental Properties Towards Novel Applications*. Springer Series in Materials Science, vol. 120 (Springer, Berlin, 2010)
[10K1] C. Klingshirn et al., Phys. Status Solidi B **247**, 1424 (2010)
[10K2] C. Klingshirn, B.K. Meyer, A. Waag, A. Hoffmann, J. Geurts, *Zinc Oxide: From Fundamental Properties towards Novel Applications*. Springer Series in Materials Science, vol. 120 (Springer, Berlin, 2010)
[10M1] B.K. Meyer in *Zinc Oxide: From Fundamental Properties Towards Novel Applications*. Springer Series in Materials Science, vol. 120 (Springer, Berlin, 2010)
[10W1] W.-J. Yin et al., Phys. Rev. B **82**, 045106 (2010)
[16G1] M. Grundmann, *The Physics of Semiconductors*, 3rd edn. (Springer International Publishing, Berlin, 2016)

Chapter 19
Excitations in One-Component Carrier Gases

Abstract The linear optical properties of heavily doped semiconductors are strongly influenced by the plasma of free carriers present in this samples. This includes of course the (degenerate) occupation of band states as well as correlation effects like collective excitations. We will address the implications of plasmons and pair excitations in the carrier gas and the description of the related optical properties in the Drude–Lorentz model. We illustrate the optical properties related to surface plasmons and of plasmon-phonon mixed states both, in bulk and low-dimensional semiconductors. We close with a discussion of prominent correlation effects like the Burstein–Moss shift and the Fermi-edge singularity.

19.1 Plasmons and Pair Excitations

The excitons presented in Chap. 20 are the energetically lowest elementary excitations of the electronic system of an ideal semiconductor (or insulator). However, if a large density of free electrons (or holes) is present in a semiconductor due to doping (Sect. 18.1) other elementary excitations appear in the electronic system. These collective excitations called *plasmons* are also well known in metals or in ionized gases.

We consider in the following a semiconductor which contains a large number of electrons (say 10^{17} upto $10^{19}\,\mathrm{cm}^{-3}$). Analogous results are found for holes.

The gas of free electrons can perform collective oscillations relative to the positive background of ionized donors. We consider in Fig. 19.1a the three-dimensional situation. A displacement of the electron gas of density n by an amount Δx produces a *surface charge density* ρ_s

$$\rho_\mathrm{s} = ne\Delta x \tag{19.1a}$$

and an electric field, according to (2.1a)

$$E_x = \frac{ne\Delta x}{\varepsilon\varepsilon_0} \tag{19.1b}$$

provided we can neglect boundary effects in the y and z directions.

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_19

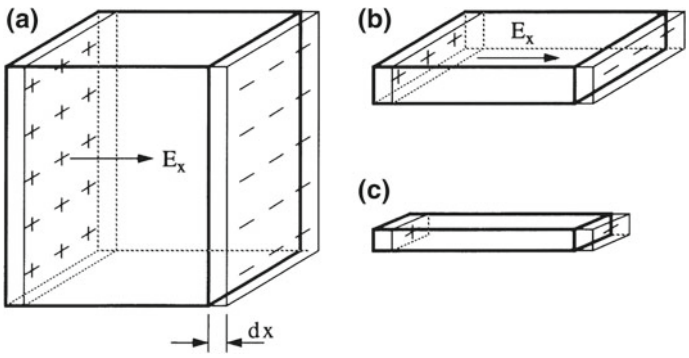


Fig. 19.1 Schematic drawing of a displacement of the electron system in a plasma oscillation in a three- (**a**), two- (**b**) and one-dimensional semiconductor (**c**)

This electric field acts back on the electrons, leading to an equation of motion

$$eE_x = \frac{e^2 n \Delta x}{\varepsilon\varepsilon_0} = m_e \frac{\partial^2}{\partial t^2}(\Delta x)\,. \tag{19.1c}$$

The solution of (19.1c) is an harmonic oscillation with frequency

$$\omega_{\mathrm{Pl}}^0 = \left(\frac{e^2 n}{m_e \varepsilon\varepsilon_0}\right)^{1/2} = \omega_{\mathrm{L}}\,, \quad \omega_{\mathrm{T}} = 0\,. \tag{19.1d}$$

The value of the dielectric function ε which enters (19.1a)–(19.1d) is the static dielectric constant ε_s provided $\omega_{\mathrm{Pl}}^0 \ll \omega_{\mathrm{L0}}$. What happens if this condition does not hold will be discussed below in Sect.19.3.

The oscillations described by (19.1a)–(19.1d) are known as *plasma oscillations.* Their quantization leads to new quasi-particles which obey Bose statistics and which are called plasmons. In gases, including the electron gas considered here, only longitudinal oscillations can propagate since gases have a non-vanishing compression but no shear stiffness (see also Sects. 7.1 and 7.3 or 8.6).

Consequently the value of ω_{Pl}^0 in (19.1d) gives the longitudinal eigenfrequency for $\lambda \to \infty$ or $\boldsymbol{k} \to 0$. The transverse eigenfrequency is zero. Proceeding to shorter wavelengths we find a weak parabolic dispersion for the plasmons [97B1, 80R1, 88R1, 02D1]

$$\omega_{\mathrm{Pl}}(\boldsymbol{k}) = \omega_{\mathrm{Pl}}^0(1 + ak^2 + \cdots) \tag{19.2}$$

shown in Fig. 19.2a.

For the densities mentioned above, $\hbar\omega_{\mathrm{Pl}}^0$, is situated in the range 10–100 meV for typical semiconductors, i.e., in the (F)IR. This situation is different in metals, where the plasma frequency is usually situated in the VIS or UV part of the spectrum and causes the high reflectivity of this class of materials which extends from the IR up to ω_{Pl}.

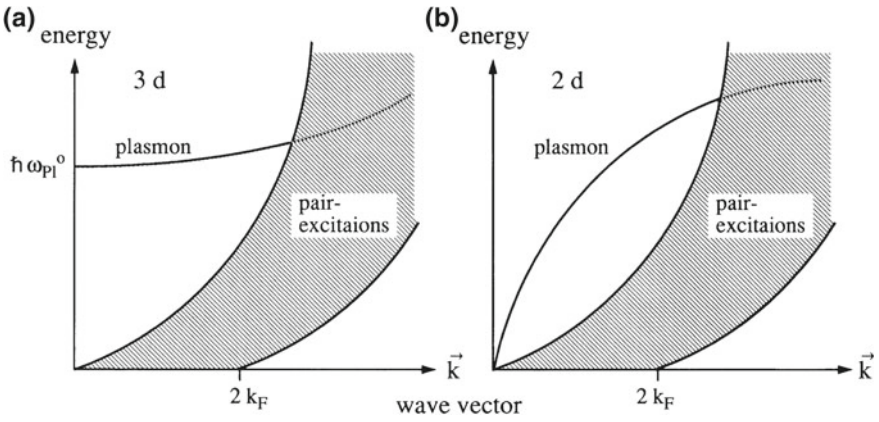


Fig. 19.2 The *plasmon dispersion* and the range of *two-particle excitations* in a three- (**a**) and a two-dimensional system (**b**)

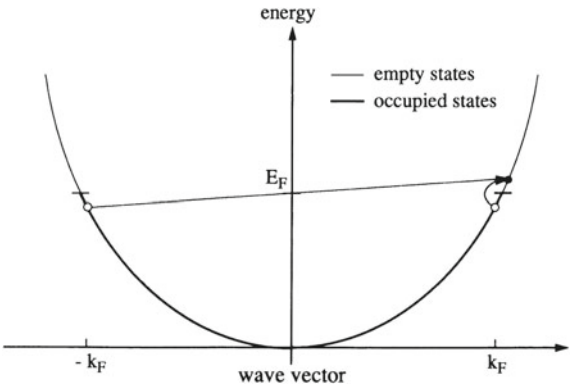


Fig. 19.3 Two-particle excitations within one band with degenerate population

For large $\boldsymbol{k}$-vectors, the plasmon modes are strongly damped because they coincide with the continuum of one-particle (or, as shown below more precisely, two-particle) intra-band excitations. These excitations are shown in Fig. 19.3 where we give the dispersion relation of the conduction band filled up to the Fermi energy E_F by a degenerate electron gas and, for simplicity, $T = 0$ K. We can produce excitations in this *Fermi sea* of electrons by lifting an electron from a state below E_F into a state above, actually by simultaneously creating a hole in the Fermi sea, below E_F, resulting in a two particle or *pair excitation*.

In contrast to the pair excitations in semiconductors and insulators across the gap, this two-particle state occurs in one band. Additionally the Coulomb interaction is substantially screened by the free carriers. We will come back to this aspect in connection with electron–hole plasmas in the second volume of this book. The excitation energies range from zero to values given by the width of the band, i.e., several eV. Small excitation energies can be created for all wave vectors between zero and $2k_F$ if the $\boldsymbol{k}$-space is at least two dimensional. For finite excitation energies the shaded range in Fig. 19.2 is accessible.

For strictly two- and one-dimensional systems, the restoring electric field $\boldsymbol{E}$ is not constant but decreases for long wavelengths as λ^{-1} or λ^{-2}, respectively. As a consequence the dispersion of plasmons starts for $k = 0$ at zero energy as shown in Fig. 19.2b, but plasmons in MQW or SL show rather the 3D dispersion relation, if the electrons are displaced in phase for many adjacent wells. For details see [67S1, 82H1, 83P1, 86H1, 89H1, 90M1, 91E1, 98P1, 99A1, 99F1, 00J1, 01O1].

The dielectric function of plasmons can be deduced in different ways, here for bulk materials. We can either use the dielectric function of a Lorentz oscillator but setting $\omega_0 = \omega_T = 0$ resulting in

$$\varepsilon\,(\omega) = \varepsilon_b + \frac{ne^2/\varepsilon_0 m}{-\omega^2 - i\omega\gamma} \tag{19.3}$$

where n is now the density of free carriers and $\gamma = 1/\tau$. Equation 19.3 again reproduces for vanishing damping the longitudinal plasma frequency of (19.1d) by setting $\varepsilon(\omega_L) = 0$.

The so-called *Drude–Lorentz model* considers bound and unbound charges separately. For bound charges the Lorentz-oscillator model of Chap. 7 is used while the equation of motion reads for free carriers for an electric field in the x-direction

$$m\ddot{x} + \frac{m}{\tau\dot{x}} = F_x = eE_x = eE_x^0 e^{-i\omega t} \tag{19.4a}$$

where $\tau = 1/\gamma$ is the relaxation time, i.e., the time between collisions resulting in a frequency-dependent electrical conductivity $\sigma(\omega)$

$$\sigma(\omega) = \sigma_0 \frac{1}{1 - i\omega\tau} \tag{19.4b}$$

and to a total dielectric function

$$\varepsilon_{\text{total}}(\omega) = \varepsilon_{\text{Lorentz}}(\omega) + \chi_{\text{free carrier}}(\omega)\,. \tag{19.4c}$$

With the help of Maxwell's equations

$$\nabla \times \boldsymbol{H} = \boldsymbol{j} + \dot{\boldsymbol{D}} = \sigma\boldsymbol{E} + \varepsilon_0\varepsilon_{\text{Lorentz}}\dot{\boldsymbol{E}}\,, \tag{19.4d}$$

we obtain for $\chi_{\text{free carrier}}(\omega)$

$$\chi_{\text{free carrier}}(\omega) = \frac{i}{\varepsilon_0\omega}\,\sigma(\omega) \tag{19.4e}$$

and for the total dielectric function

$$\varepsilon_{\text{total}}(\omega) = \varepsilon_{\text{Lorentz}}(\omega) + \frac{i}{\varepsilon_0\omega}\,\sigma(\omega) = \varepsilon_{\text{Lorentz}}(\omega) + \frac{(\omega_{Pl}^0)^2}{-\omega^2 - i\omega/\tau}\,. \tag{19.4f}$$

Since the (unscreened) plasma frequency is $ne^2/\varepsilon_0 m$, (19.4) reproduces (19.3).

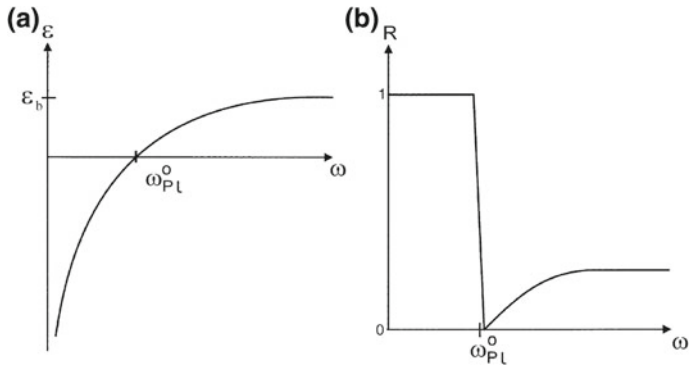


Fig. 19.4 The dielectric function of a plasmon resonance (**a**) and the resulting reflection spectrum (**b**) for vanishing damping

The $\boldsymbol{k}$-dependence of ω_{Pl}, or in other words the spatial dispersion of this resonance (see Sect. 8.4.1), can be neglected in a good approximation for bulk materials but is crucial for systems of lower dimensionality as seen from Fig. 19.2b.

If all Lorentz oscillators are situated at frequencies much higher (lower) than ω_{Pl}, their influence on the plasmon resonance can be approximated by their static (background) dielectric constant ε_{s} (ε_{b}).

We show in Fig. 19.4 the dielectric function of a plasmon resonance for such a situation for vanishing damping and the resulting reflection spectrum (compare with Figs. 7.3 and 7.6 for $\omega_0 = 0$).

In this way we can incorporate a finite conductivity into a total dielectric function. Evidently, the real (dissipative) part of the conductivity contributes to the imaginary part of $\varepsilon(\omega)$ and vice versa. In some disciplines it is also common to include all contributions in the total conductivity. But there is not only a semantic problem to address a complex conductivity of phonons or excitons, which even do not carry a net electric charge. We recommend to use the term conductivity in the context of moving carriers.

For experimental observation of plasmons as collective excitations of the carriers high doping or high optical excitation are required. Partly filled band do not usually exist in pure semiconductors in thermodynamic equilibrium, since the probability of thermal excitation across the forbidden gap is almost negligible, except for narrow gap semiconductors or for semi-metals. Under high excitation (i.e., under non-equilibrium conditions) an electron–hole plasma can be formed. It consists of electrons and holes which can acquire thermalized distributions of electrons and hole described by Fermi functions and quasi-Fermi energies in their respective bands (for details see volume two of this book). In this case the reduced mass of electron and hole enters in (19.1).

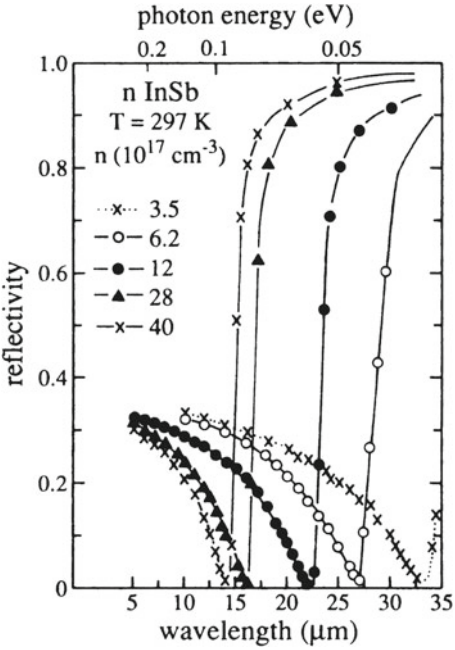


Fig. 19.5 The reflection spectrum in the vicinity of the plasmon resonance in doped InSb samples (reprinted with permission from [57S1] ©1957 American Physical Society http://doi.org/10.1007/BF00242326)

Under *high doping* and in thermodynamic equilibrium the carrier densities are given by the mass-action law (16.26)

$$np = n_{\mathrm{i}}^2(T) = N_{\mathrm{eff}}^{\mathrm{e}} N_{\mathrm{eff}}^{\mathrm{h}}\, \mathrm{e}^{-E_{\mathrm{g}}/k_{\mathrm{B}}T} \tag{19.5}$$

i.e., there is a high density either of electrons or of holes, and the data of the majority carriers have to be used in the calculation of the plasma frequency ω_{Pl} according to (19.1).

In Fig. 19.5 we show the IR reflection spectra of InSb samples with different n-doping. In agreement with our statements that the transverse eigenfrequency of plasmons is zero, we see reflectivity close to one from zero up to the plasma frequency, which is the longitudinal eigenfrequency. The reflection minimum corresponds to the frequency at which the refractive index of the upper polariton branch is unity. Above the minimum R reaches a value determined by the background dielectric constant of the plasma resonance. The shift of the reflection minimum with increasing doping reflects the $n^{1/2}$ dependence of the plasmon frequency.

19.2 Surface Plasmons

Surface plasmon modes exist in agreement with the statements in Sect. 8.6 at the frequency for which $\varepsilon(\omega) = -1$. Quantized (surface) plasmon modes exist in small metal spheres or colloids. They are, e.g., responsible for the beautiful color of gold-

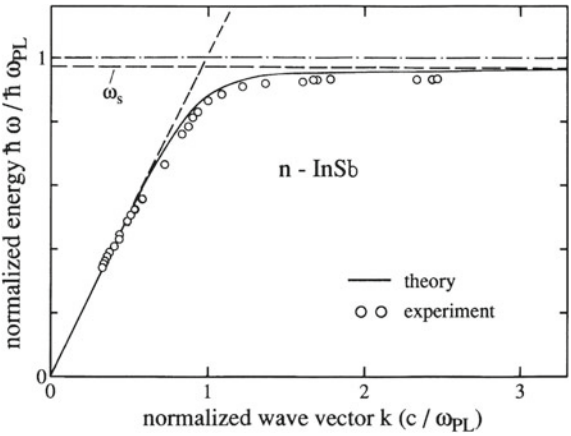


Fig. 19.6 The dispersion of the surface plasmon mode in *n*-doped InSb (reprinted with permission from [71M1] ©1971 American Physical Society http://doi.org/10.1103/PhysRevLett.27.95)

ruby glasses. We do not go further into the discussion of plasmons in metallic nanostructures, but refer the reader to a few references on this topic [09L1, 10E1, 10F1, 10G2, 10L1], which present research trends in this field including negative- and zero-index materials (see also Sect. 2.3).

Adapting Fig. 8.14 to the situation $\omega_T = 0$ we expect, for the case of plasmons, that surface plasmon modes exist in the whole frequency range between zero and ω_{Pl} (Fig. 19.6). These surface plasmons can be investigated either by attenuated total reflection (see experimental technique 21.1.5) or by the following method which has been used to obtain the data in Fig. 19.6. If a grating with spacing Λ is engraved in the surface, then the component of the wave vector parallel to the surface and normal to the grooves of the grating is conserved only modulo reciprocal vectors of this surface grating (see Sect. 5.1.3 and compare to (9.7)), i.e., modulo

$$G_{\parallel} = \frac{2\pi}{\Lambda} m;\ m = 0, \pm 1, \pm 2 \ldots . \tag{19.6}$$

Adding such $G_{\parallel}$ values to the wave vectors of the incident light beam allows coupling to the surface polariton modes for $m \neq 0$. By varying the frequency and the angle of the incident light beam and thus ω and $(\boldsymbol{k}_{\parallel} + \boldsymbol{G}_{\parallel})$ independently, it is possible to measure the dispersion of the surface plasmon polariton.

By changing the doping concentration, the current in forward direction in a *p-n* junction, or the pump power in the case of an electron–hole plasma it is possible to deliberately vary the carrier concentration in a semiconductor and thus the plasmon frequency.

19.3 Plasmon-Phonon Mixed States

The plasmon energy increases with the square-root of the carrier density. What happens when ω_{Pl}^0 approaches the energies of the optical phonons is shown in Fig. 19.7. Plasmons and phonons interact with each other due to their electric fields. This

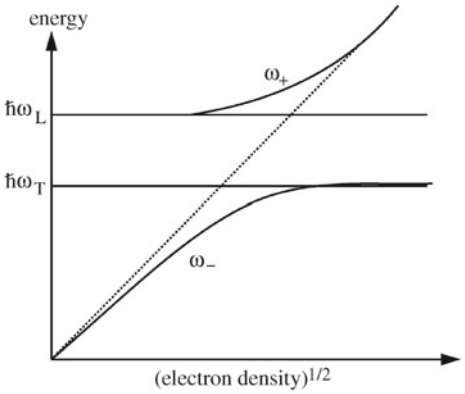


Fig. 19.7 The eigenenergies of the plasmon–phonon mixed state as a function of the square-root of the carrier density n

interaction results in another example of the quantum-mechanical non-crossing rule (Chap. 8) here as a function of the carrier density n. The plasmon frequency bends over to the transverse optical phonon branch and reappears above the longitudinal one. These two new branches are usually known as ω_- and ω_+ modes of this *plasmon-phonon mixed state*, respectively [00N1, 02N1].

This behavior can also be understood classically in the frame of linear optics by just adding the susceptibilities of a plasmon resonance and (for simplicity) of one single Lorentz oscillator, again neglecting spatial dispersion

$$\varepsilon(\omega) = \varepsilon_\mathrm{b} + \frac{f}{\omega_0^2 - \omega^2 - \mathrm{i}\omega\gamma} + \frac{(\omega_\mathrm{Pl}^0)^2}{-\omega^2 - \mathrm{i}\omega/\tau}\,. \tag{19.7}$$

The quanta of the collective excitations described by the branches ω_- and ω_+ in Fig. 19.7 correspond to plasmon-phonon mixed states [97G1, 02N1, 10K1]

We show in Fig. 19.8 the contribution of both parts separately and their sum for vanishing damping, which explains the appearance of the ω_- mode for situations when $\omega_\mathrm{Pl}^0 < \omega_0$. Care has to be taken to introduce ε_b from (19.3) and (19.4) in (19.5) only once. For plasmon frequencies considerably above those of longitudinal optical phonons, the background dielectric constant ε_b has to be used in (19.1d). For more details see. e.g. [97B1].

Depending on selection rules, plasmon-phonon mixed states may be observed experimentally in Raman scattering and/or in IR reflection spectroscopy. We give here examples for both cases. In Fig. 19.9 the longitudinal modes $\hbar\omega_-$ and $\hbar\omega_+$ of Fig. 19.7 have been detected in an electron–hole plasma as a function of electron–hole pair density confirming nicely the here developed concept.

Other examples for the observation of plasmon-phonon mixed states under optical excitation are, e.g., [02N1, 05H1]. In [05H1] it was even possible to measure the formation of the mixed state directly in the time domain. The loss function (see Fig. 12.1d) shows before and at the arrival of a 10 fs pulse on a InP sample the usual peak at the position of the longitudinal optical phonon frequency. After the arrival of

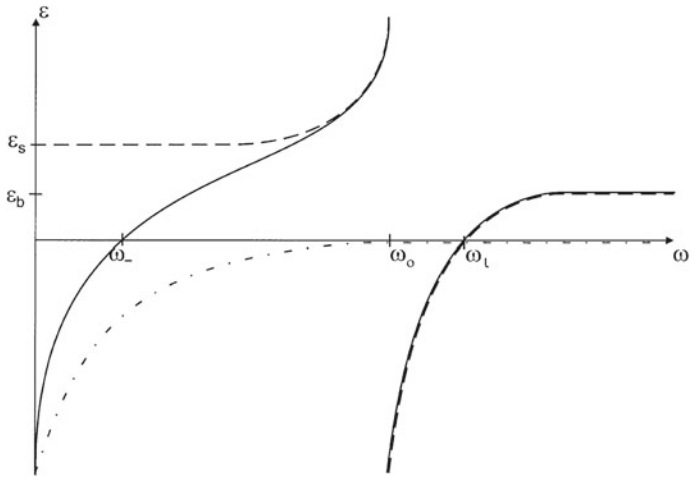


Fig. 19.8 The (undamped) dielectric functions of a Lorentz oscillator (- - -) of a plasma resonance (— · — · —·) and of the sum of both (——) (schematic)

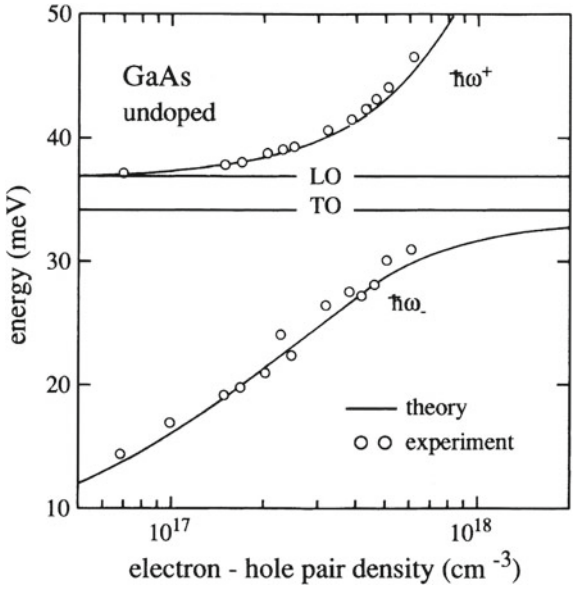


Fig. 19.9 The dependence of the plasmon–phonon mixed mode on the electron–hole pair density in highly photon-excited GaAs (reprinted with permission from [84N1] ©1984 Elsevier)

the pump pulse, two new peaks develop over a time interval of a few 10 fs. They are situated at the frequencies ω_- (just below the transverse optical phonon frequency) and at ω_+. Compare with Figs. 19.9 and 19.10c.

In Fig. 19.10a, b we show reflection spectra of ZnO:Ga with various n-type doping levels and in Fig. 19.10c the resulting transverse and longitudinal eigenenergies. One can nicely see that in Fig. 19.10a, the plasma resonance occurs below 15 meV, but in

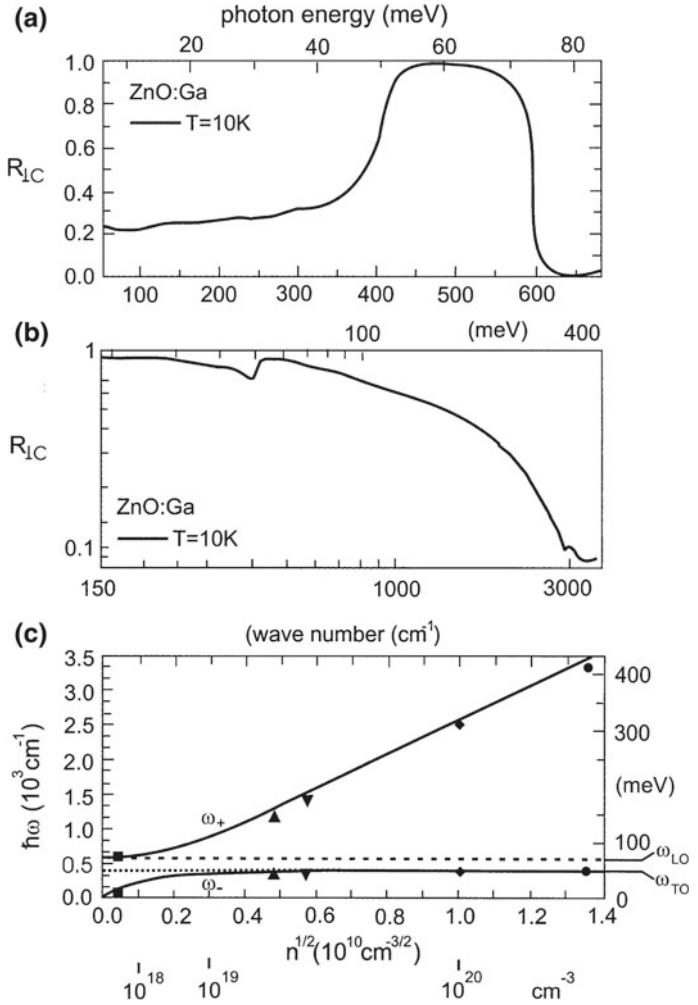


Fig. 19.10 Reflection spectra of n-doped ZnO:Ga samples with two different doping levels (note the different abscissas) (**a**, **b**) [00G1] and the various transverse and longitudinal eigenenergies of the plasmon–phonon mixed states as a function of carrier density (**c**) (reprinted with permission from [97G1] ©1997 Elsevier

Fig. 19.10b the density of $2.3 \cdot 10^{19}$ cm^{-3} gives $\hbar\omega_{\mathrm{Pl}}$ around 400 meV. The alternating stop-bands starting at the various transverse eigenfrequencies ($\hbar\omega_{\mathrm{T,Pl}} = 0$; $\hbar\omega_{\mathrm{TO}}$) and the reflection minima, occurring very close to the longitudinal eigenfrequencies $\hbar\omega_-$ and $\hbar\omega_+$, respectively, can be nicely seen [97G1, 10G1].

In simple metals the plasma frequency occurs in the 5–15 eV range, i.e., they show high reflectivity ($R \lesssim 1$) over the whole IR and visible spectral range up to the UV. The color of some metals like Cu, Au or the dip in the reflectivity of Ag in

the near UV are due to a superposition of the plasmon resonance and band-to-band transitions in a similar way as the plasmon–phonon interaction in semiconductors in the IR [81N1].

Finally it should be mentioned that there are also *valence-band plasmons* connected with collective excitations of the electron system of a filled valence band. Their eigenfrequencies are situated at energies much larger than E_g. Therefore they are not further considered here. For details of this aspect see [85E1].

19.4 Plasmons in Low-Dimensional Semiconductors

As already explained above the plasma frequency starts at $\boldsymbol{k}=0$, i.e., $\lambda=\infty$ in two- and one-dimensional systems at $\hbar\omega_{PL}=0$. See [67S1] for an early theoretical treatment of this problem. This holds, e.g., for single quantum wells and multiple quantum wells as long as the coupling between them is negligible. Besides these *intersubband plasmons* where the carriers move parallel to the QW plane, there also exist *intersubbband plasmons*. Their motion is perpendicular to the plane and one can define a transverse wavelength and $\boldsymbol{k}_\perp$. The plasma frequency of these excitations has finite values determined by the energy differences between subbands at $\boldsymbol{k}_\perp=0$ [77D1].

For layered electron gases in multiple quantum well or superlattice samples a coupled behavior of the plasmons occurs. The intraband plasmon has a dispersion $\omega_{PL}(\boldsymbol{k}_\parallel, \boldsymbol{k}_\perp)$ given, e.g., by [82O1]

$$\omega_{PL}=\left\{\frac{2\pi ne^2}{\varepsilon_M m^*_{eff}}k_\parallel\frac{\sin k_\parallel d}{\cosh k_\parallel d-\cos k_\perp d}\right\}^{1/2} \tag{19.8}$$

with n being the two-dimensional carrier density in every quantum well, m^* the effective mass of the carriers, d the distance between quantum wells and ε_M the dielectric constant of the barriers.

This equation includes the limiting cases of 2D plasmons for large separation between the wells ($\boldsymbol{k}_\parallel d\gg 1$) showing the characteristic square-root dependence, as well as the 3D case ($\boldsymbol{k}_\parallel d\ll 1$) when all the layered plasmas oscillate in phase $\boldsymbol{k}_\perp=0$). In the long-wavelength limit for the in-plane motion and for $\boldsymbol{k}_\perp\neq 0$) the induced electric fields in different layers cancel and the dispersion (19.8) reduces to a linear behavior in $\boldsymbol{k}_\parallel$.

This situation is shown in Fig. 19.11 for two different modulation-doped GaAs/$Al_{1-y}Ga_yAs$ samples. The data points have been deduced from angle-resolved Raman scattering, which allows variation of $\boldsymbol{k}_\parallel$. The solid lines have been calculated with (19.8) and the dashed lines are linear approximations.

The broadened band of the coupled plasmon modes can split into individual modes for a small number of coupled layers [86P1, 89F1].

For more details on plasmonic resonances and interactions see, e.g., [86F1, 89Z1, 93S1] or [06S1].

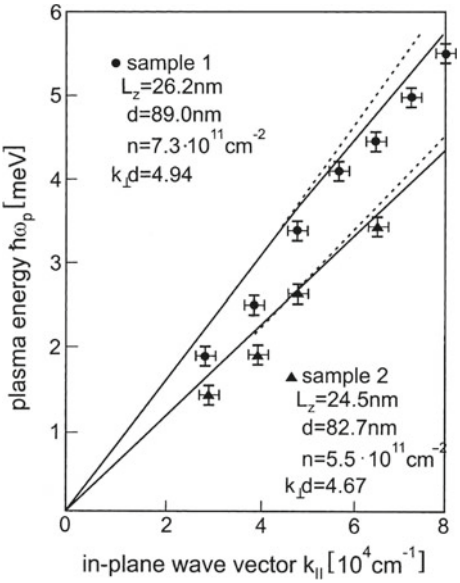


Fig. 19.11 The in-plane dispersion relation of plasmons in two different modulation-doped GaAs/$Al_{1-y}Ga_yAs$ samples (reprinted with permission from [82O1] ©1982 American Physical Society http://doi.org/10.1103/PhysRevB.25.7867)

Intensely studied are also plasmons in small nano crystals, quantum dots or anti-dot arrays. Since these structures are frequently made from metal colloids, they do not fall into the scope of this book. However, in some cases anti-dot arrays, i.e. arrays of sub-wavelength apertures (or grating slits) are manufactured in doped Si [04J1, 04R1, 05A1] or other semiconductors [06D1] and a transmission is observed exceeding by far the fraction of the surface formed by the apertures. This phenomenon results from surface plasmon polaritons, which are excited on one side of the anti-dot or grating array, are transmitted through the openings in the layer and radiate a light field on the other side of the sample.

Just as an addendum we mention again that the beautiful red color of gold ruby glasses results from surface plasmon excitation of gold colloids (or quantum dots) in the glass matrix [95K1].

19.5 Correlation Effects in Degenerately Doped Semiconductors

The optical properties of degenerately doped semiconductors (e.g., in quantum well structures having a 2D electron gas in the conduction band) are determined by Pauli's exclusion principle and many-body correlations in the electron gas as well as between the gas and excited holes. The resulting effects are schematically summarized in Fig. 19.12.

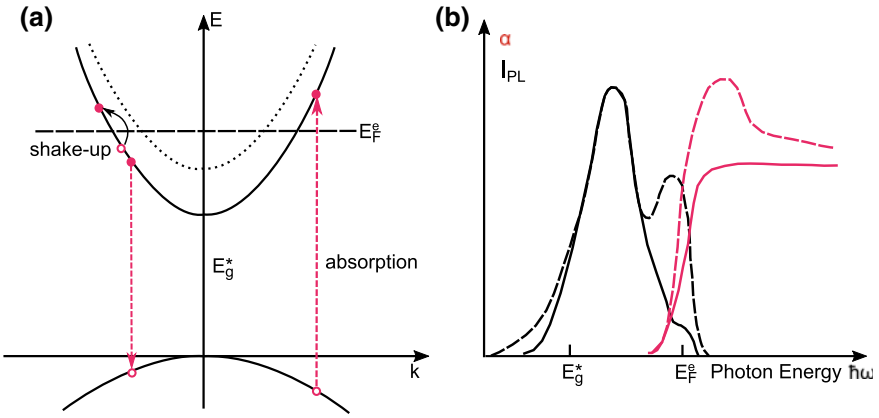


Fig. 19.12 Absorption and shake-up process in an n-doped quantum-well structure **(a)** and schematic optical spectra without (solid lines) and with (dashed lines) many-body effects **(b)**

The exclusion principle blocks optical transitions to the filled states below the Fermi level of the conduction band and leads to a blue shift of the absorption edge called *Burstein–Moss shift*. The later effect is partially compensated by the *band-edge renormalization* due to exchange and correlation effects within the one-component plasma [89C1, 92H1].

Optical spectra of doped QWs reflect the asymmetric carrier distribution of holes and electron. For a line-shape fit of PL spectra see [89C1] and for inclusion of the explicit valence-band structure see [85C1]. The spectra are further modified due to correlation effects. The screening of the positive charge of a photo-excited hole by a correlated rearrangement of the electrons close to the Fermi-edge in n-doped quantum wells leads to the so-called *Mahan exciton* [67M1]. It can in principle be observed in low-temperature optical spectra as a *Fermi-edge singularity (FES)* [87R1, 87S1]. The singularity is washed out due to the quasi-2D character of the wells [89W1]. The FES feature further broadens as a function of temperature due to the increasing carrier spread at the Fermi-edge [88L1] and by the recoil of the hole due to its finite mass [90U1, 91H1]. Since correlation effects are enhanced in low-dimensional structures the FES is typically not observed in bulk semiconductors.

Actually, optical transitions involving the states around the Fermi-edge ($\boldsymbol{k} = \boldsymbol{k}_\mathbf{F}$) in the filled conduction band require holes with equivalent momentum. Thus a (weak) FES can be readily observed in low-temperature absorption or photo-luminescence excitation (PLE) spectra of n-doped quantum wells (see e.g. [88L1, 89C1, 89K1, 91S1]. The measured enhancements of the absorption are of the order of a factor of three. In photo-luminescence the FES features are very weak [97B2]. But, the effect gets strongly enhanced when the holes are localized [87S1] or when a second unoccupied electronic band is close by [91C1, 91S1]. For a schematic depiction of optical spectra see Fig. 19.12.

The photoluminescence spectra of doped quantum wells are further modified by *shake-up processes* of the Fermi sea. During recombination of an electron with a photo-excited hole a bound excitonic state with a hole in the Fermi gas and an excited electron (symmetric in energy with respect to the Fermi energy) can be created. This leads to a lowering of the energy of the emitted photon. The result is a broadening of the low-energy side of the PL spectrum or (observed in magneto-PL) even discrete lines [90U1, 91H1, 94S1].

For FES in n- (p-)type δ-doped GaAs QWs or modulation doped asymmetric InGaAs QWs, where electron (hole) gas and holes (electrons) are spatially separated inside the well see, e.g., [90Z1, 91W1, 91W2].

The FES is also a pronounced feature in quantum wire systems (see, e.g., [93F1, 93R1, 95C1]).

For the evolution of Wannier–Mott excitons into Mahan excitons upon increasing free carrier density in n-doped bulk ZnO see [11S1].

Problems

19.1 Calculate the plasmon energy $\hbar\omega_{\mathrm{Pl}}^{0}$ for a typical three-dimensional semiconductor ($m_{\mathrm{e}} = 0.1m_0$) and $n = 10^{16}, 10^{17}$ and $10^{18}\,\mathrm{cm}^{-3}$. Compare with the eigenenergies of optical phonons.

19.2 Calculate $\hbar\omega_{\mathrm{Pl}}^{0}$ for a metal ($n \approx 10^{22} - 10^{23}\mathrm{cm}^{3}$). Using the knowledge of Chap. 7, consider which value should be taken for the dielectric "constant" ε?

19.3 What is the origin of the color of some metals like gold or copper? Remember that there are, apart from plasmons, interband-transitions in metals.

19.4 Why are radio waves in the short wave range (KW) reflected by the upper layers of the atmosphere but not ultrashort waves (UKW)?

19.5 Why are surface plasmons important for the spectral efficiency of a metal-covered diffraction grating?

19.6 Make a sketch of the dielectric function of plasmon–phonon mixed states for $\omega_{\mathrm{Pl}}^{0} > \omega_0$.

19.7 Calculate the density of electrons at which $\hbar\omega_{\mathrm{PL}} = \hbar\Omega_{\mathrm{LO}}$ for ZnO and InAs. Up to which temperatures are the electron gases degenerate?

References

[57S1] G.W. Spitzer, H.Y. Fan, Phys. Rev. **106**, 882 (1957)
[67M1] G.D. Mahan, Phys. Rev. **153**, 882 (1967)

[67S1] F. Stern, Phys. Rev. Lett. **18**, 546 (1967)
[71M1] N. Marshall, B. Fischer, H.-J. Queisser, Phys. Rev. Lett. **27**, 95 (1971)
[77D1] D.A. Dahl, L.J. Sham, Phys. Rev. B **16**, 651 (1977)
[80R1] H. Raether, *Excitation of Plasmons and Interband Transitions by Electrons*, Springer Tracts in Modern Physics, vol. 88 (Springer, Berlin, 1980)
[81N1] U. Nowak, W. Richter, G. Sachs, Phys. Status Solidi (b) **108**, 131 (1981)
[82H1] R. Höpfel et al., Surf. Sci. **113**, 118 (1982)
[82O1] D. Olego et al., Phys. Rev. B **25**, 7867 (1982)
[83P1] A. Pinczuk, J.M. Worlock, Physica **117/118B**, 637 (1983)
[84N1] H. Nather, L.G. Quagliano, Solid State Commun. **50**, 75 (1984)
[85C1] Y.-C. Chang, G.D. Sanders, Phys. Rev. B **32**, 5521 (1985)
[85E1] E. Egri, Phys. Rep. **119**, 363 (1985)
[86H1] D. Heitmann, Surf. Sci. **170**, 332 (1986)
[86F1] G. Fasol et al., Phys. Rev. Lett. **56**, 2517 (1986)
[86P1] A. Pinczuk, M.G. Lamont, A.C. Gossard, Phys. Rev. Lett. **56**, 2092 (1986)
[87R1] A.E. Ruckenstein, S. Schmitt-Rink, Phys. Rev. B **35**, 7551 (1987)
[87S1] M.S. Skolnick et al., Phys. Rev. Lett. **58**, 2130 (1987)
[88L1] G. Livescu et al., IEEE J. Quant. Electron. **24**, 1677 (1988)
[88R1] H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*, Springer Tracts in Modern Physics, vol. 111 (Springer, Berlin, 1988)
[89C1] R. Cingolani, W. Stolz, K. Ploog, Phys. Rev. B **40**, 2950 (1989)
[89F1] G. Fasol, R.D. King-Smith, D. Richards, U. Ekenberg, N. Mestres, K. Ploog, Phys. Rev. B **39**, 12695 (1989)
[89H1] M. Helm et al., Phys. Rev. Lett. **63**, 74 (1989)
[89K1] H. Kalt, K. Leo, R. Cingolani, K. Ploog, Phys. Rev. **40**, 12017 (1989)
[89Z1] T. Zettler et al., Phys. Rev. B **39**, 3931 (1989)
[89W1] J.-W. Wu, Phys. Rev. B **39**, 7992 (1989)
[90M1] U. Merkt, Festkörperprobleme/Adv. Solid State Phys. **30**, 70 (1990)
[90U1] T. Uenoyama, L.J. Sham, Phys. Rev. Lett. **65**, 1048 (1990)
[90Z1] Y.-H. Zhang, D.-S. Jiang, R. Cingolani, K. Ploog, Appl. Phys. Lett. **56**, 2195 (1990)
[91C1] W. Chen et al., Phys. Rev. B **43**, 14738 (1991)
[91E1] Th Egeler, Festkörperprobleme/Adv. Solid State Phys. **31**, 315 (1991)
[91H1] P. Hawrylak, Phys. Rev. B **44**, 3821 (1991)
[91S1] M.S. Skolnick et al., Phys. Rev. B **43**, 7354 (1991)
[91W1] J. Wagner, A. Fischer, K. Ploog, Appl. Phys. Lett. **59**, 428 (1991)
[91W2] J. Wagner, A. Ruiz, K. Ploog, Phys. Rev. B **43**, 12134 (1991)
[92H1] S. Haake et al., Phys. Rev. B **45**, 1736 (1992)
[93F1] M. Fritze, A.V. Nurmikko, P. Hawrilak, Phys. Rev. B **48**, 4960 (1993)
[93R1] F.J. Rodriguez, C. Tejedor, Phys. Rev. B **47**, 1506 (1993)
[93S1] L.C.Ó. Súilleabháin et al., Solid State Commun. **87**, 517 (1993)
[94S1] M.S. Skolnick et al., Solid State Electron. **37**, 825 (1994)
[95C1] J.M. Calleja et al., Phys. Rev. B **51**, 4285 (1995)
[95K1] U. Kreibig, R. Vollmer, *Optical Properties of Metal Clusters*, Springer Series in Material Sciences, vol. 25 (Springer, Berlin, 1995)
[97B1] R. von Baltz, *Spectroscopy and Dynamics of Collective Excitations in Solids (1995)*. NATO ASI Series B, vol. 356 (Plenum Press, New York, 1997)
[97B2] S.A. Brown, J.F. Young, Z. Wasilewski, P.T. Coleridge, Phys. Rev. B **56**, 3937 (1997)
[97G1] M. Göppert et al., J. Lumin. **72–74**, 430 (1997)
[98P1] G.A. Prinz, Science **282**, 1660 (1998)
[99A1] D.D. Awschalom, J.M. Kikkawa, Phys. Today **52**(6), 33 (1999)
[99F1] R. Fiederling et al., Nature **402**, 6783 (1999)
[00G1] M. Göppert, Dissertation, Universität Karlsruhe (TH) (2000)
[00N1] M. Nagai, M. Kuwata-Gonokami, Phys. Status Solidi B **221**, 261 (2000), J. Lumin. **100**, 233 (2002)

[00J1] C. Joachim, J.K. Gimzewski, A. Aviram, Nature **408**, 541 (2000)
[01O1] M. Oestreich et al., Festkörperprobleme/Adv. Solid State Phys. **41**, 173 (2001)
[02N1] M. Nagai, K. Ohkawa, M. Kuwata-Gonokami, Appl. Phys. Lett. **81**, 484 (2002)
[02D1] M. Dressel, G. Grüner, *Electrodynamics of Solids, Optical Properties of Electrons in Matter* (Cambridge University Press, Cambridge, 2002)
[04J1] C. Janke et al., Phys. Rev. B **69**, 205314 (2004), Opt. Lett. **30**, 2357 (2005)
[04R1] J.G. Rivas, P.H. Bolivan, H. Kurz, Opt. Lett. **29**, 1680 (2004)
[05A1] A.K. Azad, Y. Zhao, W. Zhang, Appl. Phys. Lett. **86**, 141102 (2005)
[05H1] R. Huber et al., Phys. Rev. Lett. **94**, 027401 (2005)
[06D1] E.U. Donev et al., Phys. Rev. B **73**, 201401 (2006)
[06S1] C. Schüller, *Inelastic Light Scattering of Semiconductor Nanostructures* (Springer, Berlin, 2006)
[09L1] H. Liu et al., Nat. Mater. **8**, 758 (2009)
[10E1] T. Ergin et al., Science **327**, 337 (2010)
[10F1] N. Feth et al., Opt. Express **18**, 6545 (2010)
[10G1] J. Geurts Chap. 2 in Ref. [10K1]
[10G2] S.V. Gaponenko, *Introduction to Nanophotonics* (Cambridge Univercity Press, Cambridge, 2010)
[10K1] C. Klingshirn, B.K. Meyer, A. Waag, A. Hoffmann, J. Geurts, *Zinc Oxide: from Fundamental Properties Towards Novel Applications*. Springer Series in Materials Science, vol. 120 (Springer, Berlin, 2010)
[10L1] H. Liu et al., Phys. Rev. B **81**, 241403 (2010)
[11S1] A. Schleife et al., Phys. Rev. Lett. **107**, 236405 (2011)

Part IV
Excitons

After dealing with the properties of electrons (and holes) in the semiconductor band-structure we will now focus on electron-hole pairs. They can form bound quasi-particles that dominate the optical properties of many semiconductors in wide range of conditions. In Part IV of this book on semiconductor optics we first describe the properties of Wannier excitons like binding energy, wavefunction and dispersion relation. We also introduce the triplet exciton which is typically a dark state as well as more complex species like biexcitons and trions.

Since the linear optical properties of direct-gap semiconductors with dipole-allowed band-to-band transitions are strongly governed by excitonic effects we will review spectroscopy of these materials in large detail. We will elaborate how the excitonic resonances are observed in reflection, absorption, and luminescence. Interpretation of these spectra requires knowledge on how they are affected by the polariton nature of the exciton. Spectroscopy in momentum space reveals the dispersion of the exciton polariton. Also excitons related to dipole-forbidden band-to-band transitions (like in Cu_2O), in indirect-gap semiconductors, in insulators and in organic materials will be addressed. Next, excitons bound to defects, which often dominate luminescence spectra, and excitons localized in disordered semiconductors will be discussed.

Excitons are even more prominent in low-dimensional semiconductor structures. We start with the implications of confinement in quantum wells on the exciton states and its polariton properties and how this is reflected in optical spectra. We then consider excitons in various types of superlattices. Excitons dominate the optical spectra of mono-layer semiconductors like transitionmetal dichalcogenides (TMDs) which receive particular attention due to the locking of spin and valley degrees of freedom. We will present a short introduction into the current status of research. After a summary of excitonic properties in quantum wires and rods we will review the optical properties of quantum dots (QDs). Starting from spectroscopy of QD ensembles we will focus on the single-dot spectroscopy of the excitonic finestructure in selfassembled quantum dots (SAQDs).

The next chapter is devoted to the influence of external and internal fields on the properties of excitons and the related optical properties. We start with the properties of magneto-excitons in non-magnetic as well as in semimagnetic bulk semiconductors followed by magnetic fields applied in various geometries to low-dimensional

structures. In a similar way we treat the implications of electric fields and strain fields on semiconductor optics.

Throughout this part of the book we particularly highlight some relevant experimental techniques to study the optical properties related to excitons. These techniques are the integrating sphere and the determination of the luminescence yield, two-photon absorption (TPA) and hyper-Raman scattering (HRS), attenuated total reflection (ATR), photoluminescence excitation spectroscopy (PLE), micro-photolumi-nes-cence (μ-PL) and scanning near-field microscopy (SNOM).

The final chapter of this part summarizes in an exemplary way the linear optical properties of semiconductors in the spectral range from infrared to ultraviolet light.

Chapter 20
Excitonic Quasi-particles

Abstract In this chapter we illustrate the quasi-particle properties of excitons, which are the bound states of electron–hole pair excitations in semiconductors and insulators. We will develop the concept of Wannier excitons to describe the wavefunction and dispersion relation of these excitations. We also introduce triplet excitons which are the (mostly dark) partners of the singlet ones. Then we discuss modifications of the simple exciton model due the phonon-related polarization of the crystalline lattice and introduce the exciton finestructure due to various exchange interactions. We close with an introduction to excitonic molecules (biexcitons) and charged excitons (trions).

20.1 Electron–Hole Pair Excitations

In the previous part of this book we defined the band structure for electrons and holes as the solutions to the $(N \pm 1)$-particle problem. Later we saw that the number of electrons in a band can be increased or decreased by exchange with donor and acceptor levels, respectively. In contrast, the number of electrons remains constant in the case of an optical excitation with photon energies in the band-gap region. However, an electron can be excited from the valence to the conduction band by absorption of a photon. In this process, we bring the system of N electrons from the ground state to an excited state. What we need for the understanding of the optical properties of the electronic system of a semiconductor, an insulator or even a metal is therefore a description of the excited states of the N-particle problem. The quanta of these excitations are called *excitons* in semiconductors and insulators.

We can look at this problem from various points of view: The ground state of the electronic system of a perfect semiconductor is a completely filled valence band and a completely empty conduction band. We can define this state as the *zero energy* or *vacuum state*. It has total momentum $\boldsymbol{K} = 0$, angular momentum $\boldsymbol{L} = 0$ and spin $\boldsymbol{S} = 0$. From this point $E = 0$, $\boldsymbol{K} = 0$ we will start later on to consider the dispersion relation of the excitons in connection with Fig. 20.1b.

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_20

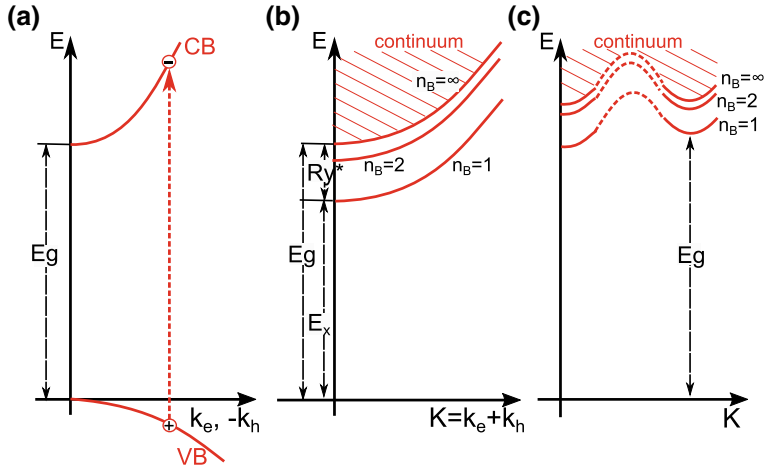


Fig. 20.1 A pair excitation in the scheme of valence and conduction band (**a**) in the exciton picture for a direct (**b**) and for an indirect-gap semiconductor (**c**)

Another point of view is the following. If we start from the above defined ground-state and excite one electron to the conduction band, we simultaneously create a hole in the valence band (Fig. 20.1a). In this sense an optical excitation is a two-particle transition. The same is true for the recombination process. An electron in the conduction band can return radiatively or non-radiatively into the valence band only if there is a free place, i.e., a hole state available. Two quasi-particles are annihilated in the recombination process.

This concept of electron–hole pair excitation is also used successfully in other disciplines of physics. If one excites, e.g., an electron in a (larger) atom or a neutron or proton in a nucleus from a deeper lying occupied state into a higher, empty one, a quantitative description concerning the transitions energies is obtained only if one takes into account both the particle in the excited state and the hole left behind. This is even true if one excites an electron in a metal from a state in the Fermi sea with an energy below the Fermi energy E_F to an empty state above.

Excitons can be described at various levels of sophistication. We present in the next sections the most simple and intuitive picture using the effective mass approximation. In the second volume of this book we discuss a more elaborate treatment of the optical properties of excitons in terms of the semiconductor Bloch equations. For other approaches and reviews on exciton research see, e.g., [57E1, 62N1, 63K1, 68L1, 77B1, 78U1, 79E1, 79R1, 79S1, 81F1, 81K1, 82R1, 85H1, 86U1, 93P1, 04K1, 06K1, 10Y1]. The concepts of *Wannier* and *Frenkel excitons*, which we will treat in the next section, were already introduced in the second half of the 1930s [31F1, 37W1].

20.2 Wannier and Frenkel Excitons

Within the effective-mass approximation, it is obvious that the Coulomb interaction between excited electron and hole (Fig. 9.1a) leads to a hydrogen-like problem with a Coulomb potential term—$e^2/(4\pi\epsilon_0\epsilon|\boldsymbol{r}_e - \boldsymbol{r}_h|)$. Indeed excitons in semiconductors form, to a good approximation, a hydrogen or positronium like series of states below the gap. For simple parabolic bands and a direct-gap semiconductor one can separate the relative motion of electron and hole and the motion of the center of mass. This leads to the *dispersion relation of excitons* shown in Fig. 20.1b:

$$E_{ex}(n_B, \boldsymbol{K}) = E_g - \mathrm{Ry}^* \frac{1}{n_B^2} + \frac{\hbar^2 \boldsymbol{K}^2}{2M} \tag{20.1a}$$

with

$$n_B = 1, 2, 3 \ldots \quad \text{principal quantum number},$$

$$\mathrm{Ry}^* = 13.6\mathrm{eV} \frac{\mu}{m_0} \frac{1}{\varepsilon^2} \quad exciton\ Rydberg\ energy, \tag{20.1b}$$

$$M = m_e + m_h \quad exciton\ translational\ mass \text{ and}$$

$$\boldsymbol{K} = \boldsymbol{k}_e + \boldsymbol{k}_h \quad exciton\ wave\ vector. \tag{20.1c}$$

For the moment, we use a capital $\boldsymbol{K}$ for the exciton wave vector to distinguish this two-particle state from the one-particle states. When we are more familiar with the exciton as a new quasi-particle we shall return to $\boldsymbol{k}$.

We also define the following quantities:

$$\mu = \frac{m_e m_h}{m_e + m_h} \quad reduced\ exciton\ mass, \tag{20.1d}$$

$$a_B^{ex} = a_B^H \varepsilon \frac{m_0}{\mu} \quad exciton\ Bohr\ radius. \tag{20.1e}$$

The radii of higher states can be considered on various levels of complexity. If one only takes into account the exponential term $\exp(Zr/na_B^H)$ in the radial part of the wave function of the hydrogen problem appearing as the envelope function in (20.4a) and defines the (excitonic) Bohr radius by the decrease of this term to $1/e$, one obtains with $Z = 1$ for excitons

$$a_B(n_B) = a_B^H \varepsilon \frac{m_0}{\mu} n_B, \tag{20.2a}$$

i.e., a linear increase with n_B.

If, on the other hand, one takes the full radial function into account, i.e., also including the factor

$$\rho^l \, L_{n_B+l}^{2l+1}(\rho) \tag{20.2b}$$

where l is the angular quantum number, $\rho = 2Z\, r/n_B \cdot a_B^H$ and $L_{n_B+l}^{2l+1}$ are the Laguerre polynomials and calculates the average distance between electron and proton or hole, respectively, one obtains [55S1]

$$\langle r(n_B)\rangle = \frac{a_B}{2}\left[3\,n_B^2 \; - \; l\,(l+1)\right] \tag{20.2c}$$

i.e., for the $n_B S$ states (i.e. $l = 0$) a quadratic dependence starting with $3a_B/2$ for $n_B = 1$.

The series of exciton states in (20.1a) has an effective Rydberg energy Ry^* modified by the reduced mass of the electron–hole pair and the dielectric constant of the medium in which these particles move. The kinetic energy term in (20.1a) involves the translational mass M and the total wave vector $\boldsymbol{K}$ of the exciton. The radius of the exciton equals the Bohr radius of the H atom again modified by ε and μ. Using the material parameters for typical semiconductors one finds

$$1\,\mathrm{meV} \le \mathrm{Ry}^* \le 200\,\mathrm{meV} \ll E_g \qquad \text{and} \tag{20.3a}$$

$$50\,\mathrm{nm} \gtrsim a_B \ge 1\,\mathrm{nm} \qquad > a_{\mathrm{lattice}}\,. \tag{20.3b}$$

This means that the excitonic Rydberg energy Ry^* is usually much smaller than the width of the forbidden gap and the Bohr radius is larger than the lattice constant. This second point is crucial. It implies that the orbits of electron and hole around their common center of mass average over many unit cells and this in turn justifies the effective-mass approximation in a self-consistent way. These excitons are called *Wannier excitons* [37W1]. They "see" the crystal as an effective medium. See also [03W1].

The one particle or $(N \pm 1)$ particle states of conduction and valence bands and the excited states of the N particle problem (or the two particle electron–hole problem) have been drawn deliberately and with good reason in different figures, namely Fig. 20.1a on one side and Fig. 20.1b, c on the other. A presentation sometimes found in text books, where the exciton dispersion and conduction and valence bands are drawn in one figure with the exciton states following the curvature of the conduction band is wrong for several reasons: one and two particle states are mixed in one diagram, the wave vector on the x-axis is simultaneously the wave vector of the electron $\boldsymbol{k}_e$ or of the hole $\boldsymbol{k}_h$ (for conduction and valence bands) or the sum of the two for the exciton (20.1c). Finally the curvature of the exciton dispersion is given by its translational mass (20.1c) and not by the electron mass.

In the above limit, excitons can usually also be described in the frame of semiconductor Bloch equations (see second volume of this text book). But finer details or corrections (see Sect. 20.3) to the simple model presented here, which are necessary to understand optical spectra, are usually not incorporated in the semiconductor Bloch equations.

It should be mentioned that in insulators like NaCl, or in organic crystals like anthracene, excitons also exist with electron–hole pair wavefunctions confined to one unit cell. These so-called *Frenkel excitons* [31F1] cannot be described in the effective-mass approximation. As a rule of thumb, one can state that in all semiconductors the inequalities (20.3a) and (20.3b) hold, so that we always deal with Wannier excitons in the following.

To get a schematic expression of the excitonic wavefunction, we use wave packets for electrons and holes $\phi_{\mathrm{e,h}}(\boldsymbol{r}_{\mathrm{e,h}})$ in the sense of the Wannier function of (14.12) or a Houston function [91C1], and obtain

$$\phi(\boldsymbol{K},n_{\mathrm{B}},l,m) = \Omega^{-1/2}\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{R}}\,\phi_{\mathrm{e}}(\boldsymbol{r}_{\mathrm{e}})\,\phi_{\mathrm{h}}(\boldsymbol{r}_{\mathrm{h}})\,\phi^{\mathrm{env}}_{n_{\mathrm{B}},l,m}(\boldsymbol{r}_{\mathrm{e}}-\boldsymbol{r}_{\mathrm{h}})\,, \tag{20.4a}$$

with the center of mass $\boldsymbol{R}$

$$\boldsymbol{R} = (m_{\mathrm{e}}\,\boldsymbol{r}_{\mathrm{e}} + m_{\mathrm{h}}\,\boldsymbol{r}_{\mathrm{h}})/(m_{\mathrm{e}}+m_{\mathrm{h}})\,, \tag{20.4b}$$

where $\Omega^{-1/2}$ is the normalization factor. The plane-wave factor describes the free propagation of a Wannier exciton through the periodic crystal similarly as for the Bloch waves of Sect. 14.2, and the hydrogen-atom-like envelope function ϕ^{env} describes the relative motion of electron and hole.

The quantum numbers l and m, with $l < n_{\mathrm{B}}$ and $-l \le m \le l$ have the same meaning as for the hydrogen atom in the limit that the angular momentum is a good quantum number in a solid. As for the H-atom, the exciton states converge for $n_{\mathrm{B}} \to \infty$ to the ionization continuum, the onset of which coincides with E_{g}.

We will see that exciton features are especially strong for regions of the electron and hole dispersions, where the group velocities of electrons and holes υ_g^e and υ_g^h are equal, e.g., zero (see, e.g., [10Y1]). This means, that in direct-gap semiconductors excitons form preferentially around $\boldsymbol{K}=0$ if the direct gap occurs at the Γ point.

In indirect-gap semiconductors exciton states form preferentially with the hole around $\boldsymbol{k}_{\mathrm{h}}=0$ and the electron in its respective minimum, as shown in Fig. 20.1c schematically e.g. for Ge where the conduction band has minima at $\boldsymbol{k}_{\mathrm{e}} = 0$ and $\boldsymbol{k}_{\mathrm{e}} \neq 0$. Electrons and holes with different group velocities still see their mutual Coulomb interaction. But, the dashed continuation of the exciton dispersion from the indirect to the direct gap is an oversimplification, among others because the states away from the band extrema are strongly damped and their binding energy varies [10Y1].

The discrete and continuum states of the excitons are resonances or oscillators which have to be incorporated into the dielectric function described in Chaps. 7 and 8. For direct semiconductors with dipole allowed band-to-band transitions, one finds an oscillator strength for excitons in discrete states with S, i.e., ($l=0$) envelope function proportional to the band-to-band dipole transition matrix element squared and to the probability of finding the electron and hole in the same unit cell. For the derivation of this relation see [57E1]. The later condition leads to the n_{B}^{-3} dependence of the oscillator strength for three-dimensional systems:

$$f_{n_B} \propto |H_{cv}^{D}|^2 \frac{1}{n_B^3} . \tag{20.5}$$

These f_{n_B} result in corresponding longitudinal-transverse splittings as shown in connection with (7.25).

Note that all excitons with finite oscillator strength can be directly created in direct-gap semiconductors (and in indirect ones by participation of a phonon $\hbar\Omega$) by absorption of a photon $\hbar\omega$ with appropriate energy, e.g., for the exciton ground state $\hbar\omega = E_g - Ry^*$ (or $\hbar\omega = E_g - Ry^* \pm \hbar\Omega$). It is not necessary, though possible, to take the detour via band-to-band excitation at or above the band gap $\hbar\omega \geq E_g$ and subsequent relaxation of the carriers into the exciton state.

Equation (20.5) actually holds for so-called *singlet excitons* with antiparallel electron and hole spin. *Triplet excitons* involve a spin flip, in their creation which significantly reduces their oscillator strength. Obviously the singlet excitons play a major role in the optical properties close to the bandgap. But also the triplet exciton states will be discussed in detail later. They play, e.g., the role of a long-lived exciton reservoir being filled by relaxation of singlet excitons after a spin flip. In the following we will give a simplified picture of these exciton states.

The crystal ground state, i.e., completely filled valence bands and completely empty conduction bands, has angular momentum $\boldsymbol{L}$, spin $\boldsymbol{S}$ and total angular momentum $\boldsymbol{J}$ all equal to zero. If an electron is optically excited from the valence band to the conduction band, e.g., by an electric dipole transition, the spin of the excited electron does not change because the electric field of the light does not act on the spin. Consequently, the simultaneously created hole has a spin opposite to the one of the excited electron and the total spin $\boldsymbol{S}$ of the electron–hole pair state is still zero (see Table 14.1). Consequently, the electron–hole pair, and likewise the exciton, is said to be in a spin singlet state. The spin $S = 1$ of the photon is accommodated by the spatial part of the band-to-band matrix element in (20.5) or by the envelope function of the exciton (see (20.4a) and [10Y1]).

If the spin flips in the transition, e.g., by interaction with the magnetic component of the light field, one ends up with a total spin $S = 1$ corresponding to spin triplet (exciton) state. The triplet state is situated energetically below the singlet state and the splitting is due to a part of the electron–hole exchange interaction, essentially the so-called *short-range* or *analytic* (for $\boldsymbol{K} \to 0$) part of the *exchange interaction*. There are names other than triplet and singlet in use for certain materials like *para* and *ortho exciton* (e.g., for the case of Cu_2O) or *bright* and *dark exciton* (e.g., in quantum dots). The last pair of names reflects the fact that triplet-exciton resonances have small oscillator strength because they are spin-flip forbidden.

Since spin and angular momentum are, strictly speaking, no good quantum numbers in a solid (see Appendix B on group theory) it is not obligatory that the triplet exciton is threefold degenerate and the singlet is not degenerate. There are cases where the triplet or para exciton is non-degenerate, as in Cu_2O, and the singlet is threefold degenerate, as in many zinc-blende type crystals like CuCl, ZnSe or GaAs. This is possibly the reason that in some older, especially French, literature the names singlet and triplet exciton states are interchanged [62N1].

Recently it has been proposed in [18B1] that triplet excitons in caesium lead halide perovskites ($CsPbX_3$) are bright. The reason is that a mixture of Pb 6s and X 4p forms s-like Bloch states of the valence band while strong spin-orbit coupling mixes the Pb 6p orbital with spin to form $J = 1/2$ conduction-band states. This results in dark singlet and bright triplet exciton states. If additionally the Rashba effect (spin-orbit coupling plus inversion asymmetry) is present, then the triplet states can even be the lowest exciton states.

The oscillator strength of the excitonic continuum states is actually influenced by the so-called *Sommerfeld enhancement* factor. We will come back to this point later, when we discuss the optical properties.

In the picture of second quantization, we can define creation operators for electrons in the conduction band and for holes in the valence band $\alpha^+_{\boldsymbol{k}_\mathrm{e}}$ and $\beta^+_{\boldsymbol{k}_\mathrm{h}}$, respectively. The combination of both gives creation operators for electron–hole pairs $\alpha^+_{\boldsymbol{k}_\mathrm{e}}\beta^+_{\boldsymbol{k}_\mathrm{h}}$. The *exciton creation operator* B^+ can be constructed via a sum over electron–hole pair operators (see [81K1], or [93H1]):

$$B^+_{\boldsymbol{K}} = \sum_{\boldsymbol{k}'_\mathrm{e}\boldsymbol{k}_\mathrm{h}} \delta[\boldsymbol{K} - (\boldsymbol{k}_\mathrm{e} + \boldsymbol{k}_\mathrm{h})]\alpha_{\boldsymbol{k}_\mathrm{e},\boldsymbol{k}_\mathrm{h}}\alpha^+_{\boldsymbol{k}_\mathrm{e}}\beta^+_{\boldsymbol{k}_\mathrm{h}} \,. \tag{20.6}$$

The expansion coefficients $a_{\boldsymbol{k}_\mathrm{e},\boldsymbol{k}_\mathrm{h}}$ correspond, in principle, to those also used in an (slightly old-fashioned) expansion into Slater determinants of the many particle problem, which contains either all valence band states $\varphi_{\boldsymbol{k}_h}(\boldsymbol{r}_{i,h})$ for the ground state or always has one line being replaced by a conduction-band state [63K1].

It can be shown that the $B^+_{\boldsymbol{K}}$ and $B_{\boldsymbol{K}}$ obey Bose commutation relations with a density-dependent correction term which increases with the number of electrons and holes contained in the volume of one exciton $4\pi(a^{\mathrm{ex}}_\mathrm{B})^3/3$ [77H1, 81K1]. This has two consequences: in thermodynamic equilibrium for low densities and not too low temperatures, the excitons can be well described by Boltzmann statistics with a chemical potential ruled by their density and temperature similar to (16.25). For higher densities they deviate more and more from ideal bosons until they end up in an *electron–hole plasma* made up entirely from fermions. This makes the creation of a *Bose–Einstein condensate* of excitons (or of biexcitons) a very complicated problem. We will treat these many-particle phenomena in the second volume of this book.

20.3 Corrections to the Simple Exciton Model

The simplified model outlined in the preceding section is, as already mentioned, adequate for non-degenerate, parabolic bands. We keep these assumptions for the moment and inspect a first group of corrections which are relevant for the parameters ε and μ entering in (20.1a)–(20.1e). Since ε is a function of ω there is always the question of which value should be used.

As long as the binding energy of the exciton $E_{\mathrm{ex}}^{\mathrm{b}}$ is small compared to the optical phonon energies and, consequently, the excitonic Bohr radius (20.1e) larger than the polaron radius (14.20d)

$$E_{\mathrm{ex}}^{\mathrm{b}} < \hbar\omega_{\mathrm{LO}}, \quad a_{\mathrm{B}} > r_{\mathrm{p}} \qquad \Rightarrow \varepsilon = \varepsilon_{\mathrm{s}}\,, \tag{20.7a}$$

we can use for ε the static value ε_{s} below the phonon resonances and the polaron masses and polaron gap. This situation is fulfilled for some semiconductors for all values of n_{B}, e.g., for GaAs where $\mathrm{Ry}^* \simeq 5\,\mathrm{meV}$ and $\hbar\omega_{\mathrm{LO}} \simeq 36\,\mathrm{meV}$.

In many other semiconductors including especially the wide-gap semiconductors (see Fig. 20.2) the inequality (20.7a) holds only for the higher states $n_{\mathrm{B}} \geq 2$, while for the ground state exciton ($n_{\mathrm{B}} = 1$) we find

$$E_{\mathrm{ex}}^{\mathrm{b}} \gtrsim \hbar\omega_{\mathrm{LO}}, \quad a_{\mathrm{B}} \simeq r_{\mathrm{p}} \qquad \Rightarrow \varepsilon_{\mathrm{s}} \geq \varepsilon \geq \varepsilon_{\mathrm{b}}\,. \tag{20.7b}$$

Examples are CdS, ZnO, CuCl or Cu_2O.

In this situation a value for ε between ε_{s} and ε_{b} seems appropriate, because the polarization of the lattice can only partly follow the motion of electron and hole. A useful approach is the so-called *Haken potential* [55H1] which interpolates between ε_{s} and ε_{b} depending on the distance between electron and hole, where r_{eh}, $r_{\mathrm{e}}^{\mathrm{p}}$ and $r_{\mathrm{p}}^{\mathrm{h}}$ are the distances between electron and hole, and the polaron radii of electron and hole, respectively:

$$\frac{1}{\varepsilon(\boldsymbol{r}_{\mathrm{e,h}})} = \frac{1}{\varepsilon_{\mathrm{b}}} - \left(\frac{1}{\varepsilon_{\mathrm{b}}} - \frac{1}{\varepsilon_{\mathrm{s}}}\right)\left(1 - \frac{\exp(-\boldsymbol{r}_{\mathrm{eh}}/\boldsymbol{r}_{\mathrm{e}}^{\mathrm{p}}) + \exp(-\boldsymbol{r}_{\mathrm{eh}}/\boldsymbol{r}_{\mathrm{h}}^{\mathrm{p}})}{2}\right). \tag{20.8}$$

The next correction concerns the effective masses. The polarization clouds of the polarons (Sect. 14.6) have different signs for electron and hole. If both particles are bound together in an exciton state fulfilling (20.7b) the polaron renormalization is partly quenched, with the consequence that values for the effective masses will lie somewhere between the polaron values and the ones for a rigid lattice. The gap seen by the exciton in the $1S$ state will likewise be situated between the two above extrema. Fortunately, the above effects tend to partly compensate each other. A transition from the polaron gap to the larger rigid lattice gap shifts the exciton energy to larger energies. A transition from ε_{s} to ε_{b} and an increase of the effective masses increases the binding energy and shift the $1S$ exciton to lower photon energies. As a consequence one finds, even for many semiconductors for which inequality (20.7b) holds, that the $1S$ exciton fits together with the higher exciton states reasonably well into the hydrogen-like series of (20.1). A good example is ZnO with $\mathrm{Ry}^* = (59 \pm 1)\,\mathrm{meV}$ and $\varepsilon_{\mathrm{s}} = 8.5$ and $\varepsilon_{\mathrm{b}} = 3.7$ [04M1]. The exciton series follows very precisely a n_{B}^{-2} law but with an effective ε around 6.4 [75H1]. We shall use this approach in the future if not stated otherwise.

In other cases like Cu_2O or CuCl, the higher states with $n_{\mathrm{B}} \geq 2$ follow a hydrogen-like n_{B}^{-2} series converging to the polaron gap, but the $1S$ exciton shows, with respect to the polaron gap, a binding energy $E_{\mathrm{ex}}^{\mathrm{b}}$, which differs from and is generally larger

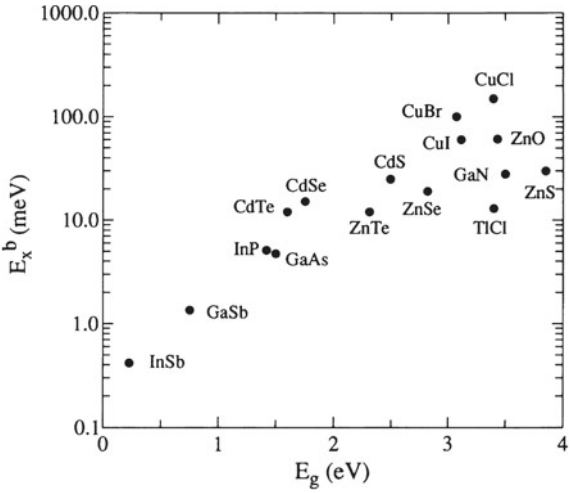


Fig. 20.2 The exciton binding energy E_{ex}^{b} as a function of the band-gap energy E_g for various direct-gap semiconductors. The data are compiled from [82L1, 93H1, 93P1]

than the excitonic Rydberg energy deduced from the higher states. We call the experimentally observed energetic distance between the $1S$ exciton and the polaron gap the *exciton binding energy* E_{ex}^{b}, in contrast to Ry^* in (20.1). This discrepancy introduces some ambiguity when comparing theoretical results with experimental data, since in theory one often normalizes energies with the excitonic Rydberg energy Ry^*, but the $1S$ exciton has a different value for E_{ex}^{b} (see also [63K2]).

There is a general trend of the material parameters m_{eff} and ε with E_g which results in an increase of the exciton binding energy with increasing E_g as shown in Fig. 20.2. This is a consequence analogous to increasing the band width or in other words decreasing effective mass with decreasing width of the gap as described, e.g., in the band structure model of Sect. 14.5.

The next complication arises from details of the band structure. If the bands are degenerate, as is the Γ_8 valence band in T_d symmetry, it is no longer possible to separate the relative and the center of mass motion—they are coupled together. Similar effects stem from k-linear terms and other sources. We get light- and heavy-hole exciton branches and splittings, e.g., between the 2S and 2P exciton states partly induced by the envelope function. An example for Cu_2O will be given in Sect. 21.2.

Furthermore it should be mentioned, without going into details, that the splitting between singlet and triplet excitons Δ_{st} and the splitting of the singlet state into a transverse and a longitudinal one Δ_{LT}, are both due to exchange interaction between electron and hole. This caused by their Coulomb interaction [73D1, 88F1, 03G1] if we consider the N-electron problem in the form of Slater's determinant where the ground state consists only of valence-band states and the excited state of a sum of determinants in each of which one valence-band state is replaced by a conduction-band state. This aspect is treated in detail in [73D1, 93H1, 93P1].

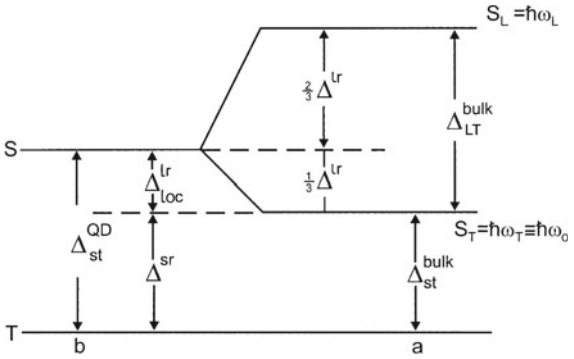


Fig. 20.3 The splitting of the 1S exciton with the notation of the various contributions to *exchange splitting* for bulk samples (right) and in a quantum dot (left). The abbreviations have the following meanings: QD: quantum dot, S: singlet, T: triplet, Δ^{lr}: long range or non-analytic (for $\boldsymbol{k} \to 0$) part of the exchange interaction, Δ^{sr}: short range or analytic part of the exchange interaction

The mentioned splittings are illustrated in Fig. 20.3. Usually the following relation holds for Wannier excitons:

$$\Delta_{st} \ll \Delta_{LT} \quad \text{with} \quad 0.1\,\text{meV} \lesssim \Delta_{LT} \lesssim 15\text{meV}. \tag{20.9}$$

For the $A\Gamma_5$-excitons in CdS or ZnO one finds, e.g., $\Delta^{CdS}_{LT} = 1.8\,\text{meV}$, $\Delta^{CdS}_{st} = 0.2\,\text{meV}$ and $\Delta^{ZnO}_{LT} = 1.2\,\text{meV}$, $\Delta^{ZnO}_{st} = 0.17\,\text{meV}$. Since the short-range interaction increases with decreasing exciton radius as a_B^{-3} the situation begins to change for 1S excitons with a value of a_B^{ex} exceeding only slightly the lattice constant, and leading thus also to the limit of the concept Wannier excitons. One finds, e.g., for CuCl or Cu_2O: $\Delta^{CuCL}_{LT} = 5.5\text{meV}$, $\Delta^{CuCl}_{st} = 2.5\text{meV}$, $\Delta^{Cu_2O}_{LT} \approx 50\,\mu\,\text{eV}$, $\Delta^{Cu_2O}_{st} = 12\,\text{meV}$ (see [85H1]). The extremely low value of Δ_{LT} in Cu_2O comes from the fact, that the band-to-band transition is parity forbidden and the 1S singlet or ortho-exciton is only allowed in quadrupole approximation (see Sect. 21.2).

The singlet–triplet splitting is enhanced by compression of the excitonic wave function in quantum wells and especially in quantum dots (see Fig. 20.3 and, e.g., [98F1, 03G1]). We will come back to this aspect in Sect. 23.5.

It should be mentioned that the excitation of an optically (dipol-) allowed exciton is accompanied by a polarization as detailed in Chap. 8 on polaritons and in volume two on (semiconductor) Bloch equations, respectively.

Excitons can also couple to phonons to form exciton-phonon bound states. This concept has been introduced in [68L1, 68T1, 72K1, 02B1].

Finally we mention that excitons can also be formed with holes in deeper valence bands. These so-called *core-excitons* are usually situated in the VUV or X-ray region of the spectrum and have a rather short lifetime. An example and further references are given in Sect. 21.1.7 and [87C1, 88K1].

20.4 Biexcitons and Trions

When introducing the concept of excitons, we stated that they can be understood on a certain level of sophistication as the analogue to the hydrogen atom or—even better—to the positronium atom. This analogy can be even pushed further. It is well known that two hydrogen atoms with opposite electron spin can bind to form a hydrogen molecule. In the same sense it has been calculated, that two positronium atoms can bind to a positronium molecule [09F1]. So it is not to surprising that two excitons can bind to form a new quasiparticle, the so-called *biexciton* or *excitonic molecule* [58L1, 58M1].

It has been found theoretically that the biexciton should form a bound state for all ratios of effective electron and hole masses as well as for any dimensionality of the sample. The *biexciton binding energy* expressed in units of the excitonic Rydberg energy starts for bulk samples for $\sigma = m_e/m_h \to 0$ at $E^b_{biex}/Ry^* \approx 0.3$ corresponding to the value of a hydrogen molecule and then drops monotonically for rising σ, reaching values of 0.027 or 0.12 for $\sigma = 1$ [72A1, 72B1, 73H1, 81K1]. It approaches for $\sigma \to \infty$ again the value of 0, since the four-particle problem is symmetric against the exchange of electrons and holes, i.e., $\sigma \to 1/\sigma$.

The dispersion relation of the biexciton is given in the simplest case by

$$E_{biex}(\boldsymbol{k}) = 2(E_g - E^b_{ex}) \; - \; E^b_{biex} \; + \; \frac{\hbar^2 \boldsymbol{k}^2}{4M_{ex}} \tag{20.10}$$

assuming that the effective mass of the biexciton is just twice that of the exciton. However, many of the complications mentioned for excitons in Sect. 20.3 also apply here.

Furthermore it has been predicted theoretically [76S1, 97S1] and verified experimentally, at least for quantum structures [95F1, 97S2], that *trions* also form bound states contributing to the luminescence below the free-exciton energy. Trions are charged excitons or biexcitons, i.e., quasiparticles consisting of two electrons and one hole or vice versa. Their observation is favored by a moderate n or p (modulation) doping of quantum well structures. Localization effects may play a role as well (see e.g. [04K2]).

As we shall see in the second volume of this book, biexcitons and trions have indeed been observed experimentally in a wide variety of bulk semiconductors with both direct and indirect band gaps.

Problems

20.1 Calculate the Rydberg energy and the Bohr radius of excitons for some important semiconductors. The required material parameters can be found in Appendix C or, e.g., in [82L1]. Compare these with the experimentally determined binding energies and lattice constants, respectively.

20.2 How many (different) exciton states can be constructed in a semiconductor with zinc-blende (T_d) structure for the principal quantum numbers $n_B = 1, 2$ and 3?

20.3 Compare the magnitude of the relative splitting between 2s and 2p states in a hydrogen atom (what are the physical reasons?) with the 2s–2p splitting of excitons.

20.4 Plot the Rydberg series of an idealized three- and two-dimensional exciton and indicate the oscillator strengths.

20.5 Calculate the (combined) density of states in the continuum of a three- and a two-dimensional exciton in the effective-mass approximation. Multiply by the corresponding Sommerfeld enhancement factor.

20.6 Find in the literature data for the binding energies of the exciton ground state and of the higher states (i.e. $n_B S$ or $n_B P$ states with $n_B \geq 2$), e.g., for GaAs, ZnO, CuCl and Cu_2O and determine for which ones the $1S$ state fits into the hydrogen series with higher states.

References

[31F1] J. Frenkel, Phys. Rev. **37**, 17, 1276 (1931), Phys. Z. Sowjetunion **9**, 158 (1936)
[37W1] G.H. Wannier, Phys. Rev. **52**, 191 (1937)
[55H1] H. Haken, Halbleiterprobleme IV, 1 (1955). Nuovo Cimento **3**, 1230 (1956)
[55S1] L.I. Schiff, *Quantum Mechanics*, 2nd edn. (Mc Graw Hill, New York, 1955)
[57E1] R.J. Elliot, Phys. Rev. **108**, 1384 (1957)
[58L1] M.A. Lampert, Phys. Rev. Lett. **1**, 450 (1958)
[58M1] S.A. Moskalenko, Opt. Spectrosc. **5**, 147 (1958)
[62N1] S. Nikitine, Prog. Semicond. **6**(233), 269 (1962)
[63K1] R.S. Knox, *Theory of Excitons*. Solid State Physics, Suppl. 5 (Academic, New York, 1963)
[63K2] C.G. Kuper, G.D. Whitfield (eds.), *Polarons and Excitons* (Plenum, New York, 1963)
[68L1] W.Y. Liang, A.D. Yoffe, Phys. Rev. Lett. **20**, 59 (1968)
[68T1] Y. Toyozawa, J. Hermanson, Phys. Rev. Lett. **21**, 1637 (1968)
[72A1] O. Akimoto, E. Hanamura, J. Phys. Soc. Jpn. **33**, 1537 (1972), Solid State Commun. **10**, 253 (1972)
[72B1] W.F. Brinkman, T.M. Rice, B. Bell, Phys. Rev. B **8**, 1570 (1972)
[72K1] L. Kalok, J. Treusch, Phys. Status Solidi A **52**, K 125 (1972)
[73D1] M.M. Denisov, V.P. Makarov, Phys. Status Solidi (b) **56**, 9 (1973)
[73H1] W.T. Huang, Phys. Status Solidi (b) **60**, 309 (1973)
[75H1] A. Haug, Festkörperprobleme XII **411** (1975)
[76S1] B. Stébé, C. Conte, Solid State Commun. **19**, 1237 (1976)
[77B1] D. Bimberg, Festkörperprobleme/Adv. Solid State Phys. **17**, 195 (1977)
[77H1] H. Haug, E. Hanamura, Phys. Rep. **33C**, 209 (1977)
[78U1] R.G. Ulbrich, C. Weisbuch, Festkörperprobleme/Adv. Solid State Phys. **18**, 217 (1978)
[79E1] K. Cho (ed.), *Excitons*, Topics in Current Physics, vol. 14 (Springer, Berlin, 1979)
[79R1] U. Rössler, Festkörperprobleme/Adv. Solid State Phys. **19**, 77 (1979)
[79S1] A. Stahl, Ch. Uihlein, Festkörperprobleme/Adv. Solid State Phys. **19**, 159 (1979)
[81F1] D. Fröhlich, Festkörperprobleme/Adv. Solid State Phys. **21**, 363 (1981)

[81K1] C. Klingshirn, H. Haug, Phys. Rep. **70**, 315 (1981)
[82L1] O. Madelung, U. Rössler (ed.), *Landolt–Börnstein*. New Series, Group III, vol. 17 a to i, 22 a and b, 41 A to D (Springer, Berlin, 1982–2001)
[82R1] E.I. Rashba, M.D. Sturge (Eds.): *Excitons*, Modern Problems in Condensed Matter Sciences, vol. 2 (North-Holland, Amsterdam, New York 1982)
[85H1] B. Hönerlage et al., Phys. Rep. **124**, 161 (1985)
[86U1] M. Ueta, H. Kazanki, K. Kobayashi, Y. Toyozawa, E. Hanamura, *Excitonic Processes in Solids*. Springer Series in Solid State Sciences, vol. 60 (Springer, Berlin, 1986)
[87C1] R.D. Carson, S.E. Schnatterly, Phys. Rev. Lett. **59**, 319 (1987)
[88F1] C.Y. Fong, I.P. Batra, S. Ciraci (eds.), *Properties of Impurity States in Superlattice Semicondcutors*. NATO ASI Series B, vol. 183 (Plenum, New York, 1988)
[88K1] M. Krause, H.-E. Gummlich, U. Becker, Phys. Rev. B **37**, 6336 (1988)
[91C1] J.N. Churchill, F.E. Holmstrom, Phys. Scr. **44**, 395 (1991)
[93H1] H. Haug, S.W. Koch, *Quantum Theory of the Optical and Electronic Properties of Semiconductors*, 2nd edn. (World Scientific, Singapore, 1993)
[93P1] N. Peyghambarian, S.W. Koch, A. Mysyrowicz, *Introduction to Semiconductor Optics* (Prentice Hall, Englewood Cliffs, 1993)
[95F1] G. Finkelstein, H. Shtrikman, I. Bar Joseph, Phys. Rev. Lett. **74**, 976 (1995), Phys. Rev. B **53**, 12593, R1709 (1996)
[97S1] B. Stébé et al., Phys. Rev. B **56**, 12454 (1997)
[97S2] A.J. Shields et al., Phys. Rev. B **55**, R1970 (1997)
[98F1] A. Franceschetti, L.W. Wang, A. Zunger, Phys. Rev. B **58**, 13367 (1998)
[02B1] M. Betz et al., Phys. Status Solidi b **231**, 181 (2002)
[03G1] O. Gogolin et al., J. Lumin.**102/103**, 414, 451 (2003)
[03W1] H-Ch. Weissker, J. Furthmüller, F. Bechstedt, Phys. Rev. B **67**, 165322 (2003)
[04K1] H. Kalt, M. Hetterich (eds.), *Optics of Semiconductors and Their Nanostructures*. Springer Series in Solid State Sciences, vol. 146 (Springer, Berlin, 2004)
[04K2] A. Klochikhin et al., Phys. Rev. B **69**, 085308 (2004)
[04M1] B.K. Meyer et al., Phys. Status Solidi b **241**, 231 (2004)
[06K1] S.W. Koch et al., Nat. Mater. **5**, 523 (2006)
[09F1] A.M. Frolov, Phys. Rev. A **80**, 014502 (2009)
[10Y1] P.Y. Yu, M. Cardona, *Fundamentals of Semiconductors*, 4th edn. (Springer, Heidelberg, 2010)
[18B1] M.A. Becker et al., Nature **553**, 189 (2018)

Chapter 21
Optical Properties of Intrinsic Excitons in Bulk Semiconductors

Abstract Phonon excitations are necessary to describe the optical properties of semiconductors and of insulators in the IR. Plasmons determine the optical properties of metals from the IR through the visible to the near UV, and in semiconductors, if present at all, they contribute along with the phonons to the IR spectra. Excitons, on the other hand, determine together with their continuum states (or the band-to-band transitions) the optical properties around and above the band gap, i.e., in the visible including the near UV and IR in the case of semiconductors and in the (V)UV for insulators. Although inorganic insulators like the alkali halides and organic ones such as anthracene have specific optical properties, many of the aspects presented in the following for excitons in semiconductors also apply to them. We will present in this chapter the intrinsic linear optical properties of excitons in bulk semiconductors starting from semiconductors with a dipole-allowed, direct band-to-band transition. They exhibit dipole-allowed excitons with the highest oscillator strength. Values of their longitudinal–transverse splitting Δ_{LT} range from 0.1 to beyond 10 meV. However, not all excitons in this group of semiconductors have high oscillator strength. Also some excitons in semiconductors with dipole-forbidden band-to-band transitions may be dipole allowed, but with considerably lower oscillator strength. We will review these cases in Sect. 21.2, ending with some comments on indirect gap materials and intra-excitonic transitions (Sect. 21.3). Again, several experimental techniques will be introduced like use of integrating spheres, spectroscopy in momentum space and attenuated total reflectance.

21.1 Excitons with Strong Oscillator Strength

In semiconductors with dipole-allowed direct band-to-band transitions, excitons couple strongly to the radiation field in the case of not too strong damping. As a consequence many optical properties can be understood quantitatively only in the strong-coupling or polariton picture. Thus we use this occasion to elucidate once more for this particular case the concepts of weak and of strong coupling to the radiation field. We will review the consequences of spatial dispersion on the optical properties and elaborate the signatures of excitons in optical spectra. The we will illustrate methods to determine the polariton dispersion with spectroscopic methods in momentum

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_21

space. We will discuss surface-exciton polaritons, excitons in organic materials and insulators and finally more exotic species like saddle-point and core excitons.

21.1.1 Exciton–Photon Coupling

In Chap. 2 we introduced the electromagnetic radiation field and in Chap. 3 the photons as its quanta. Their interaction with various excitations of the solid (phonons, plasmons, excitons ...) can be treated in perturbation theory. This is the *weak coupling* approach described in Chap. 6. Here the one-photon absorption coefficient $\alpha(\omega)$ between an initial and a final state is in first order-perturbation theory proportional to the dipole matrix element squared:

$$\alpha(\omega) \propto |\langle f \mid H_{\mathrm{D}} \mid i\rangle|^2 \, \delta(E_{\mathrm{i}} - E_{\mathrm{f}} + \hbar\omega) \,. \tag{21.1a}$$

The refractive index is obtained at this level of approximation either by a Kramers–Kronig transformation of $\alpha(\omega)$ or, away from the resonance, by second-order perturbation theory according to

$$n^2(\omega) - 1 \propto \sum_{z \neq i} \frac{\langle i|H_{\mathrm{D}}|z\rangle \, \langle z|H_{\mathrm{D}}|i\rangle}{E_z - E_{\mathrm{i}} - \hbar\omega} \tag{21.1b}$$

A photon $\hbar\omega$ creates virtually an excited intermediate state $|z\rangle$ under momentum conservation, which, after a time Δt limited by

$$\Delta E \, \Delta t = |E_2 - E_1 - \hbar\omega| \, \Delta t \lesssim \hbar \,, \tag{21.2}$$

emits again a photon which is identical to the incident one, while the electronic system returns to the initial state $|i\rangle$. The time Δt during which the energy is "stored" in the virtually excited state reduces the phase velocity of the light and thus evidently describes an $n(\omega)$ which increases when ω approaches the resonance energy $E_z - E_{\mathrm{i}}$ from below, since ΔE goes to zero and Δt can be very long, in agreement with Fig. 5.5.

In the polariton concept, on the other hand, one quantizes the mixed state of the electromagnetic radiation and the excitation of the medium, i.e., the polarization wave. We already introduced this concept in Chap. 8. Since it is a very important one we want to demonstrate it here again for the *exciton polariton*.

For readers who are not satisfied with the simple statement that the polaritons are the quanta of the mixed states of electromagnetic radiation and excitation (or polarization), we give two other approaches. The first is just a diagrammatic representation (Fig. 21.1) of what was said before: an incident photon creates an electron–hole pair, which recombines again to give a photon, and so on. The Coulomb interaction between electron and hole, which is responsible for the formation of the exciton, is represented by a virtual exchange of photons between electron and hole, i.e., by the

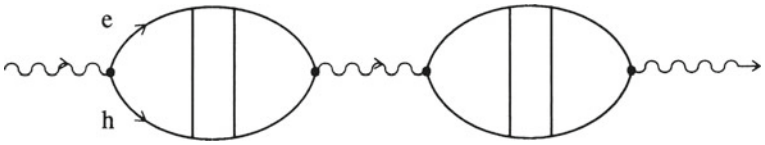


Fig. 21.1 Diagrammatic representation of an exciton polariton

vertical lines. Consequently the whole diagram of Fig. 21.1 can be considered as a representation of the exciton polariton.

In the other approach, which follows [93H1], we start with the electron and hole operators, construct from them the exciton and finally the exciton polariton (see also Sect. 20.2). We start with the creation and annihilation operators for excitons:

$$B^{+}_{v,\boldsymbol{k}}\,;\; B_{v,\boldsymbol{k}}\,. \tag{21.3}$$

The index v stands for the quantum numbers n_{B}, l, m. It can be shown that the $B_{v,\boldsymbol{k}}$ deviate from the commutator relations of ideal bosons by a term proportional to the mean number of electron–hole pairs n contained in the volume of an exciton a^{d}_{B} (see e.g., [93H1, 93P1])

$$\langle [B_{0,0}\,,\, B^{+}_{0,0}]^{-}\rangle = 1 - \mathrm{O}(n a^{d}_{\mathrm{B}})\,, \tag{21.4}$$

where d is the dimensionality of the system. The Hamiltonian of a non-interacting exciton gas is then

$$H = \sum_{v,\boldsymbol{k}} E(v,\boldsymbol{k})\, B^{+}_{v,\boldsymbol{k}} B_{v,\boldsymbol{k}}\,. \tag{21.5}$$

Using an analogous expression for the photons with the number operator $c^{+}_{\boldsymbol{k}} c_{\boldsymbol{k}}$, for the interacting system of excitons and photons and considering the leading, i.e., resonant terms around a specific resonance only, we obtain

$$H = \sum_{\boldsymbol{k}} \left[\sum_{v} E_{v\boldsymbol{k}}\, B^{+}_{v\boldsymbol{k}} B_{v\boldsymbol{k}} + \hbar\omega_{\boldsymbol{k}}\, c^{+}_{\boldsymbol{k}} c_{\boldsymbol{k}} - \mathrm{i}\,\hbar \sum_{v} g_{v\boldsymbol{k}}\, (B^{+}_{v\boldsymbol{k}}\, c_{\boldsymbol{k}} - h.c.) \right]. \tag{21.6}$$

The coupling coefficients $g_{v\boldsymbol{k}}$ contain the transition matrix elements as in (21.1).

If we consider the third term on the right-hand side of (21.6) as a perturbation, we are back once more to the weak coupling limit.

The polariton concept is obtained if we diagonalize the whole Hamiltonian (21.6) by a suitable linear combination of the $B_{v\boldsymbol{k}}$ and the $c_{\boldsymbol{k}}$, leading to the *polariton operator* $P_{\boldsymbol{k}}$:

$$P_k = u_{v\boldsymbol{k}}\, B_{v,\boldsymbol{k}} - v_{\boldsymbol{k}}\, c_{\boldsymbol{k}} \tag{21.7}$$

$$\text{with}\quad |u_{v\boldsymbol{k}}|^2 + |v_{\boldsymbol{k}}|^2 = 1\,. \tag{21.8}$$

The $u_{v\boldsymbol{k}}$ give the exciton-like character of the polaritons. They are close to one around the exciton energy $E_{v,0}$ and decrease on the upper and lower polariton branches with decreasing energetic distance from the resonance. The $v_{\boldsymbol{k}}$ give the photon-like part and, according to (21.8), show the opposite behavior [93H1]. As a rule of thumb we can state that the polariton wavefunction contains considerable exciton-like parts over energies $|\hbar\omega - E_{\mathrm{ex}}| \lesssim 10\Delta_{\mathrm{LT}}$.

It is interesting to note that the dispersion relation that we obtain from this approach is identical to the one obtained from the set of classical coupled oscillators treated in Chaps. 7 and 8.

21.1.2 Consequences of Spatial Dispersion

In contrast to that of optical phonons, the $\boldsymbol{k}$-dependence of the exciton energy is significant. For $\boldsymbol{k}$-vectors in the transition region from the photon-like to the exciton-like part of the dispersion relation, the so-called *bottle neck*, the kinetic energy term in (20.1a) becomes comparable to the longitudinal-transverse splitting Δ_{LT}. The consequences of *spatial dispersion* have already been outlined in Sect. 8.4, so that we can restrict ourselves to just recalling them here. For all frequencies there is at least one propagating mode. This fact reduces the reflectivity in the reststrahlen band to values below 1, even in the case of negligible damping. For frequencies above ω_{L} there are several propagating modes, and below it there is at least one propagating mode and one or more evanescent ones. This situation is not covered by the boundary conditions deduced from Maxwell's equations and *additional boundary conditions* have to be introduced containing the information about what fraction of the energy transmitted through the interface travels on which polariton branch. Since this branching ratio is ω-dependent and since the imaginary parts of the various branches differ, the decay of the intensity into the depth of the sample can be non-exponential. This means the "effective" absorption coefficient can be thickness dependent. Furthermore the absorption spectrum looks more complex (Fig. 8.12) than without spatial dispersion (Fig. 7.4).

The abc which have been introduced by Pekar and by Hopfield [58H1, 62H1, 62P1, 63H1, 64M1] assume that the excitonic part of the polarization at the surface vanishes, or its derivative normal to the surface, or a linear combinations of both (see Sect. 8.4 and the literature given there). In [98H1] a way out of this abc problem has been shown. For even more recent approaches see [00T1, 01S1]. Furthermore one can assume that excitons do not "leak out" of the semiconductor into vacuum and that there should consequently be an *exciton-free surface layer* (*dead layer*), the optical properties of which are described by ε_{b} and which has a minimum thickness of the excitonic Bohr radius. Electric fields, which occur often normal to surfaces can ionize, i.e., destroy the exciton and lead to an increase in the thickness of the dead layer. The problem which thus has to be solved to calculate a reflection spectrum is shown in Fig. 21.2. An incident beam passes first the dead layer in which multiple reflection occurs and then enters the semiconductor in which several modes can be

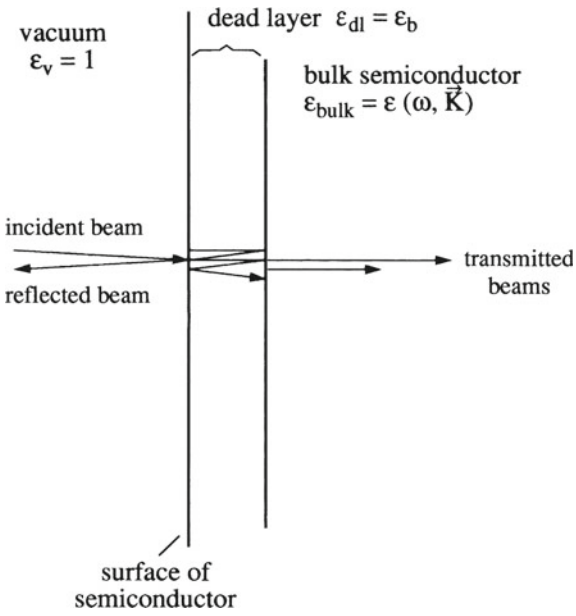


Fig. 21.2 The problem of reflection for a semiconductor in the vicinity of an exciton resonance for normal incidence including multiple reflection in a dead layer and two propagating modes due to spatial dispersion

excited. Sometimes scientists apply even more complex models assuming, e.g., that the damping and/or the eigenfrequencies are depth dependent [79E1, 81L1, 82S1]. Finally, in [04S1] it has been suggested that the exciton free layer may be absent at interfaces between similar materials.

The formulas to calculate optical spectra including abcs are rather complex and we do not give them here but refer the reader to [75L1, 78S1, 79E1, 80B1, 81L1, 82B1, 82S1, 84R1, 85H1, 93K1, 95B1]. Instead we give in the next section examples of reflection, transmission, and luminescence spectra of the exciton polariton in bulk semiconductors with direct, dipole-allowed band-to-band transitions.

21.1.3 Signature of Excitons in Optical Spectra

21.1.3.1 Reflection Spectra

In order to illustrate the appearance of excitonic signatures in optical spectroscopy we start with the example of the II-VI compound CdS. In Fig. 21.3a we show the bandstructure of CdS and in (b) the dispersion of the exciton polariton for the orientation $\boldsymbol{k} \perp \boldsymbol{c}$, $\boldsymbol{E} \perp \boldsymbol{c}$ of the polariton wave. We include the $n_B = 1$ excitons involving a hole either in the A or in the B valence band. The A exciton is a rather simple resonance for this orientation, comparable to our model system in Sect. 8.2. The k-linear term of the B valence band (Sect. 15.2) mixes the singlet and triplet states for $\boldsymbol{k}_\perp \neq 0$ (see Sect. 21.2.1.1) and gives rise to an additional polariton branch.

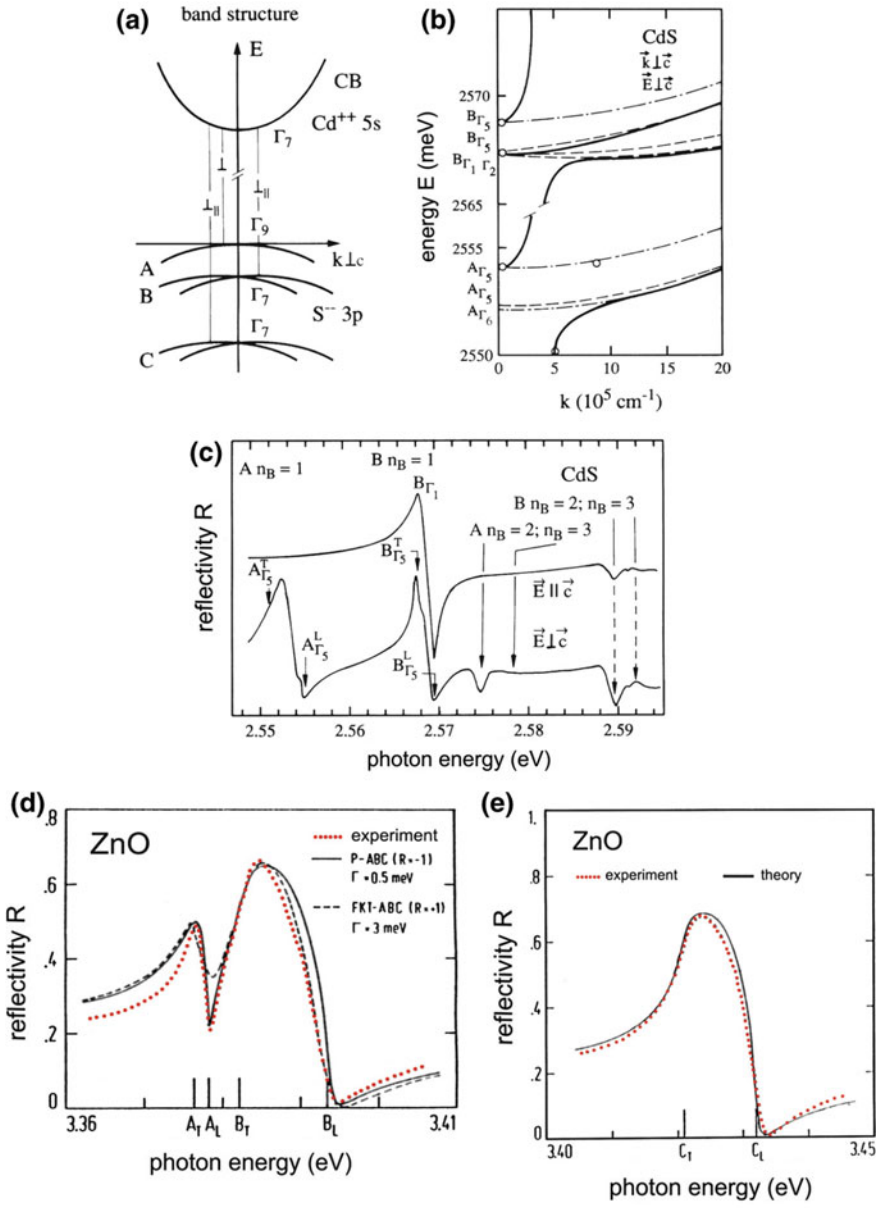


Fig. 21.3 The bandstructure of CdS around the Γ-point (**a**); the dispersion of the $n_B = 1$ A and B exciton polariton resonances (**b** reprinted with permission from [85H1]) ©1985 Elsevier); the reflection spectra for the polarizations $\boldsymbol{E} \perp \boldsymbol{c}$ and $\boldsymbol{E} \parallel \boldsymbol{c}$ (**c** reprinted with permission from [82B1] ©1982 American Physical Society https://doi.org/10.1103/PhysRevB.25.7413) and the reflection spectra for ZnO for the $n_B = 1$ exciton resonances in both orthogonal polarizations (**d**, **e**) [78H1]

Figure 21.3c finally gives the reflection spectra of the two resonances for $\boldsymbol{E} \perp \boldsymbol{c}$ and $\boldsymbol{E} \parallel \boldsymbol{c}$ and of some higher states ($n_B \geq 2$) (see also [73H1, 75L1, 78H2]). The combination of the $\Gamma_1(S$-)envelope function for $n_B = 1$ with the symmetries of the electron Γ_7 and the holes ($A\Gamma_9$, $B\Gamma_7$) gives excitons of symmetries $A\Gamma_5$ and $A\Gamma_6$ and $B\Gamma_1$, $B\Gamma_2$, $B\Gamma_5$ as explained in more detail in Appendix B. The Γ_6 and Γ_2 states are triplets which couple only weakly to the radiation field since they are spin-flip and dipole-forbidden and do not show up in reflection (see Sect. 21.2.1.1). Γ_5 and Γ_1, states couple to the radiation field for the orientations $\boldsymbol{E} \perp \boldsymbol{c}$ and $\boldsymbol{E} \parallel \boldsymbol{c}$, respectively. These selection rules show up clearly in the reflection spectra, the $n_B = 1$A excitons being seen only in $\boldsymbol{E} \perp \boldsymbol{c}$.

A fit to the reflection spectra (not shown here) using spatial dispersion, an exciton-free layer and some abc coincides with experiments within a few percent.

The $A\Gamma_5$ resonance is quite simple, as already mentioned. R remains significantly below 1 due to spatial dispersion as predicted above. For smaller oscillator strength or longitudinal-transverse splitting, the maximum almost disappears and only a narrow dip close to the longitudinal eigenfrequency remains, as shown in Fig. 21.4 for ZnTe.

The small spike around $A\Gamma_5^L$ in Fig. 21.5c is partly caused by the onset of the UPB but mainly by the dead layer. An increase of its thickness increases the importance of the spike due to multiple reflections (Fig. 21.4) and may even lead to an inversion of the usual reflection spectrum, i.e., to a dip at low energies and a maximum above. A set of calculated spectra showing this phenomenon is given in Fig. 21.5.

The $B\Gamma_1$ exciton resonance is again a simple one, but the $B\Gamma_5$ has a small dip stemming from the additional polariton branch shown in Fig. 21.3b, which at this energy reaches exactly $n = 1$.

While the A and $B\Gamma_5$ excitons have in CdS for $\boldsymbol{E} \perp \boldsymbol{c}$ roughly equal oscillator strength and Δ_{LT} [85H1], the situation changes for close-lying resonances as occurs for ZnO [85H1]. In this case the longitudinal-transverse splitting Δ_{LT} of the $A\Gamma_5$ exciton is reduced and that of the $B\Gamma_5$ increased as discussed already in Sect. 7.5. Only the sum of $\Delta_{LTA\Gamma_5}$ and $\Delta_{LTB\Gamma_5}$ has to be constant and equal to $\Delta_{LTC\Gamma_1}$ [60H1]. To illustrate this statement, Fig. 21.3d, e show a close-up of the $n_B = 1$ exciton reflection spectra for the two orthogonal polarizations $\boldsymbol{E}$ perpendicular and parallel to $\boldsymbol{c}$. Note the different selection rules for CdS and ZnO, resulting from the inverted valence band ordering in ZnO [81R1, 10K1, 10K2].

At higher energies we see $n_B \geq 2$ exciton states which split into several sublevels due to the various L and/or L_z values of the envelope function. The reflection signal of these higher states decreases due to the n_B^{-3} dependence of the oscillator strength [57E1]. In the band-to-band transition region the reflection spectra are usually flat and structureless. The C-excitons expected for CdS from Fig. 21.3a are situated around 2.61 eV and are off the scale of Fig. 21.3b, c. These resonances are washed out even at low temperature because the C-exciton is situated in the continuum of the A- and B-excitons and thus has a rather short phase relaxation time T_2, i.e., strong damping. For the case of ZnO the $C\Gamma_A$ exciton is shown in Fig. 21.3e.

With increasing lattice temperature the exciton resonances are broadened due to increasing scattering with phonons. Sometimes they are hardly visible at RT as shown in Fig. 21.6 for CdS. Similar data are also found for ZnO e.g., in [80B2]. A

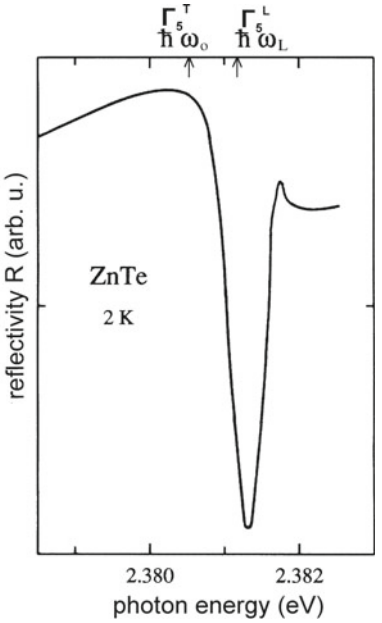


Fig. 21.4 A reflection spectrum for ZnTe (adapted with permission from [83M1] ©1983 Springer Nature

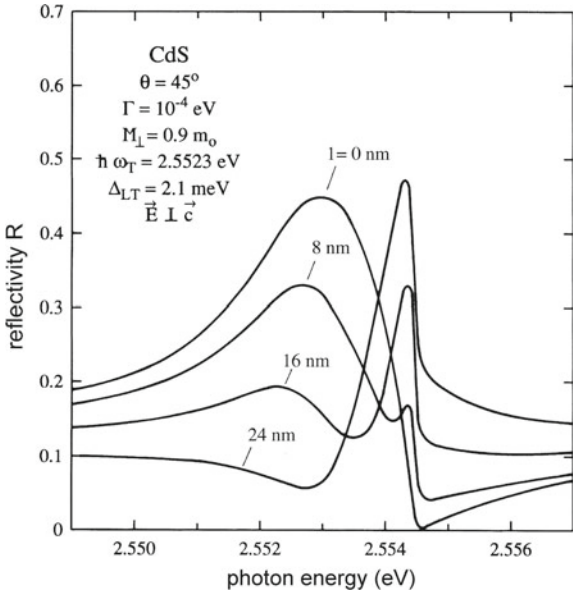


Fig. 21.5 A set of calculated exciton reflection spectra of CdS for 45° incidence and for various thicknesses of the exciton-free layer with all other parameters kept constant (according to [82R1])

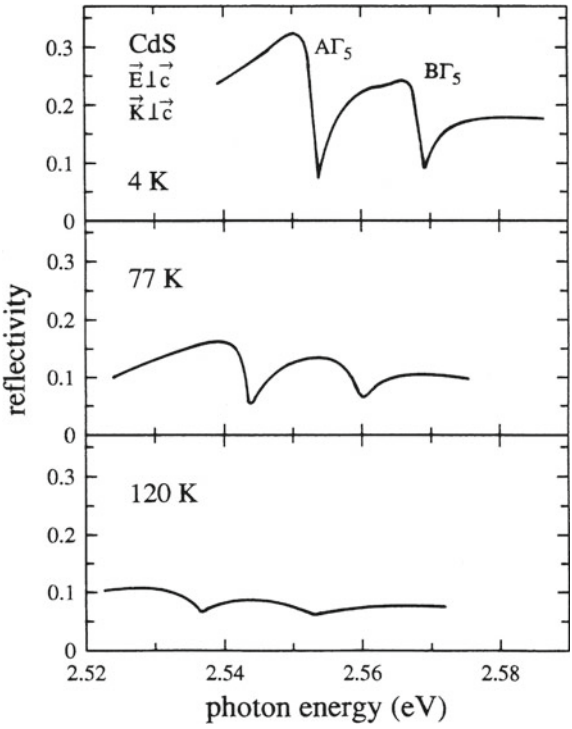


Fig. 21.6 A set of reflection spectra of the A- and B-exciton resonances of CdS for various lattice temperatures (adapted with permission from [80B2] ©2006 Wiley and Sons)

similar washing out of the exciton resonance can occur even at low temperatures in samples with high impurity content and/or lower crystalline quality or for increasing excitation density.

We will come back to the temperature dependence of the damping of excitons [90R1] in connection with luminescence spectroscopy below. For the temperature dependence of the band gap itself see e.g., [67V1, 99P1, 07K1, 10K1, 10K2].

Many more examples of reflection spectra III–V-, II–VI- and I–VII compounds showing excitonic signatures are found e.g., in the references above, in [77V1, 79V1, 80B1, 80D1, 81K1, 85H1, 86U1], in [74B1] for AgI, or in the data collection [82L1].

To summarize, we can state that the low-temperature reflection spectra of direct-gap semiconductors are determined around the gap by exciton polaritons. The longitudinal eigenenergy can be reasonably well deduced from the reflection minimum which corresponds to $n = 1$ on the UPB and is therefore situated only slightly ($<$1 meV) above $\hbar\omega_L$. The transverse eigenenergy, the oscillator strength, the damping, and the effective mass of the exciton can be extracted only with the help of a rather complicated line-shape analysis.

Some complications can arise in uniaxial, hexagonal materials like CdS or ZnO due to the occurrence of *mixed-mode polaritons* in these birefringent materials. It can be shown from group-theoretical considerations that the Γ_5-excitons are the resonances for the ordinary beam (Sect. 5.4.1) because the polarization $\boldsymbol{E} \perp \boldsymbol{c}$ can

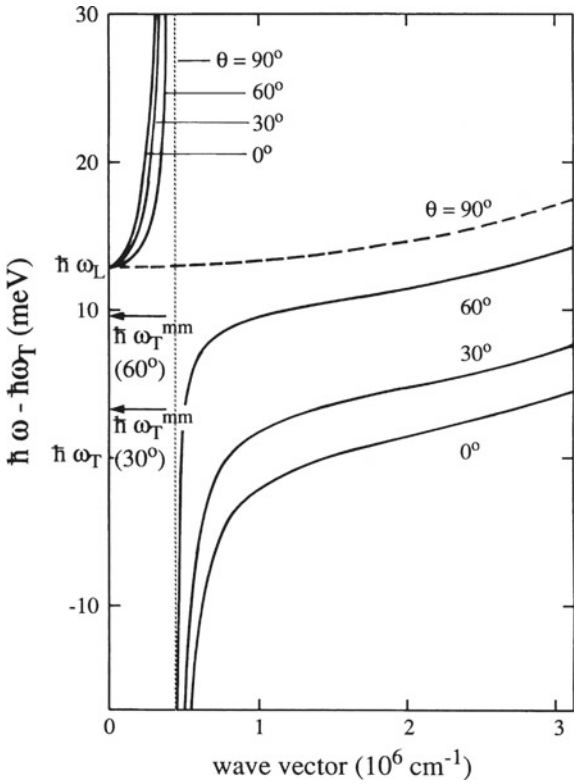


Fig. 21.7 Dispersion of the exciton polariton in a uniaxial material for various angles between $\boldsymbol{k}$ and $\boldsymbol{c}$. Data for the CΓ_1 exciton polariton in ZnO (According to [78H1])

be realized for all angles between $\boldsymbol{k}$ and $\boldsymbol{c}$. The Γ_1 resonances have dipoles oriented parallel to $\boldsymbol{c}$. As a consequence they have maximum coupling to the radiation field for $\boldsymbol{E} \parallel \boldsymbol{c}$ and $\boldsymbol{k} \perp \boldsymbol{c}$, but develop as extraordinary or mixed-mode polaritons to the longitudinal state if the angle $\angle(\boldsymbol{k}, \boldsymbol{c})$ is continuously changed from $\boldsymbol{k} \perp \boldsymbol{c}$ to $\boldsymbol{k} \parallel \boldsymbol{c}$. In Fig. 21.7 we show the dispersion of the CΓ_1 exciton polariton in ZnO, which has an oscillator strength varying according to

$$\Delta_{\mathrm{LT}} = \Delta_{\mathrm{LT}}^{0} \sin^2 \angle(\boldsymbol{k}, \boldsymbol{c}) \,. \tag{21.9}$$

The complimentary angle dependence holds for the A and BΓ_5 excitons for the orientation of $\boldsymbol{E}$ in the main section, i.e., for extraordinary beams.

Experimental data for mixed-mode polaritons can be found in [72W1, 75L1]. The small reflection spike seen at B$\Gamma_{5\mathrm{L}}$ in [03C1] might similarly well be due to a mixed-mode polariton due to the finite angle of aperture of the incident beam. For a more exhaustive discussion of the reflection spectra of ZnO see [07K1, 10K1, 10K2].

21.1.3.2 Absorption Spectra

Absorption spectra of $n_B = 1$ polariton resonances are usually difficult to measure quantitatively, since in the resonance region the effective absorption (or rather the extinction) coefficient reaches values in the range of $10^4\,\mathrm{cm}^{-1} \lesssim \alpha_{\mathrm{eff}} \lesssim 10^6\,\mathrm{cm}^{-1}$. As a consequence for samples with $d \gtrsim 1\,\mu\mathrm{m}$, the transmitted light intensity goes to "zero", i.e., to values comparable to or smaller than the stray light of the spectrometer.

We show in Fig. 21.8a an absorption spectrum of a ZnO layer at room temperature (RT). The layer is of poly-crystalline nature, produced by evaporating Zn on a substrate of quartz glass and subsequent oxidation [59H1]. The crystallites have their $\boldsymbol{c}$ axis oriented perpendicular to the substrate. The peak at 3.3 eV is due to the close lying and thermally broadened $A\Gamma_5$ and $B\Gamma_5$ exciton resonances [68L1, 75L1, 78H1, 80B1, 81L1, 82B1, 85H1, 03C1]. This is one of the first (unrecognized at that time) experimental observations of a Wannier exciton (Sect. 20.2) in absorption.

In Fig. 21.8b experimental transmission spectra show A and $B\Gamma_5$ resonances of a thin CdS platelet-type sample. There are Fabry–Perot modes (see Sect. 3.3) in the transparent region due to the natural reflectivity of the surfaces of the as-grown platelet. A detailed analysis of $\alpha_{\mathrm{eff}}(\omega)$ in the resonance region is possible only for sample thicknesses of the order of 1 μm, such that $\alpha d \lesssim 4$. Such samples cannot usually be produced by grinding and polishing of thicker samples, since these processes introduce so much damage to the lattice that the broadening γ of the resonance becomes too large (i.e., $\hbar\gamma > \Delta_{\mathrm{LT}}$).

Sometimes epitaxial layers can be used for this type of investigation, but care has to be taken that these layers do not contain (inhomogeneous) strain since transmission measurements integrate over the whole sample thickness. Fortunately some semiconductors, such as ZnO, CdS or CdSe, tend to grow as thin, single crystal platelets. The absorption spectrum of Fig. 21.8b stems from such a sample. Figure 21.8c shows a calculated transmission spectrum of the $A\Gamma_5$ $n_B = 1$ resonance. The similarity to the A_1 resonance in Fig. 21.8b and the deviations from the simple case without spatial dispersion of Fig. 7.5 or 12.1 are evident.

In Fig. 21.8d we give an overview of the absorption features of thin GaAs samples. GaAs has, in comparison to CdS or ZnO, a much lower oscillator strength due to the larger value of a_B (see Sect. 20.2): $\Delta_{\mathrm{LT}}(\mathrm{GaAs}) \approx 0.1$ meV. Therefore it is easier to measure the absorption of the $n_B = 1$ exciton. We can see in this figure similarly to CdS in Fig. 21.8b also the $n_B = 2$ and 3 levels with S-envelope function. Even higher states ($n_B > 3$) merge with the continuum. The decrease of the oscillator strength with n_B^{-3} [57E1] is at least qualitatively confirmed by this spectrum.

In Fig. 21.8e and f we finally show absorption spectra of a thin ZnO platelet for the two main polarizations and for 4.2 and 295 K. One sees again the Rydberg series at 4.2 K up to $n_B = 3$ followed by higher energy structures involving LO phonons. If the coupling to LO phonons is strong enough, exciton-phonon bound states may develop [68T1, 72K1, 82J1]. Similar structures have been observed in the Cu-halides [66R1, 72B1] or in the photoluminescence excitation spectra of CdTe and CdS [75H1] or in PbI_2 [94W1].

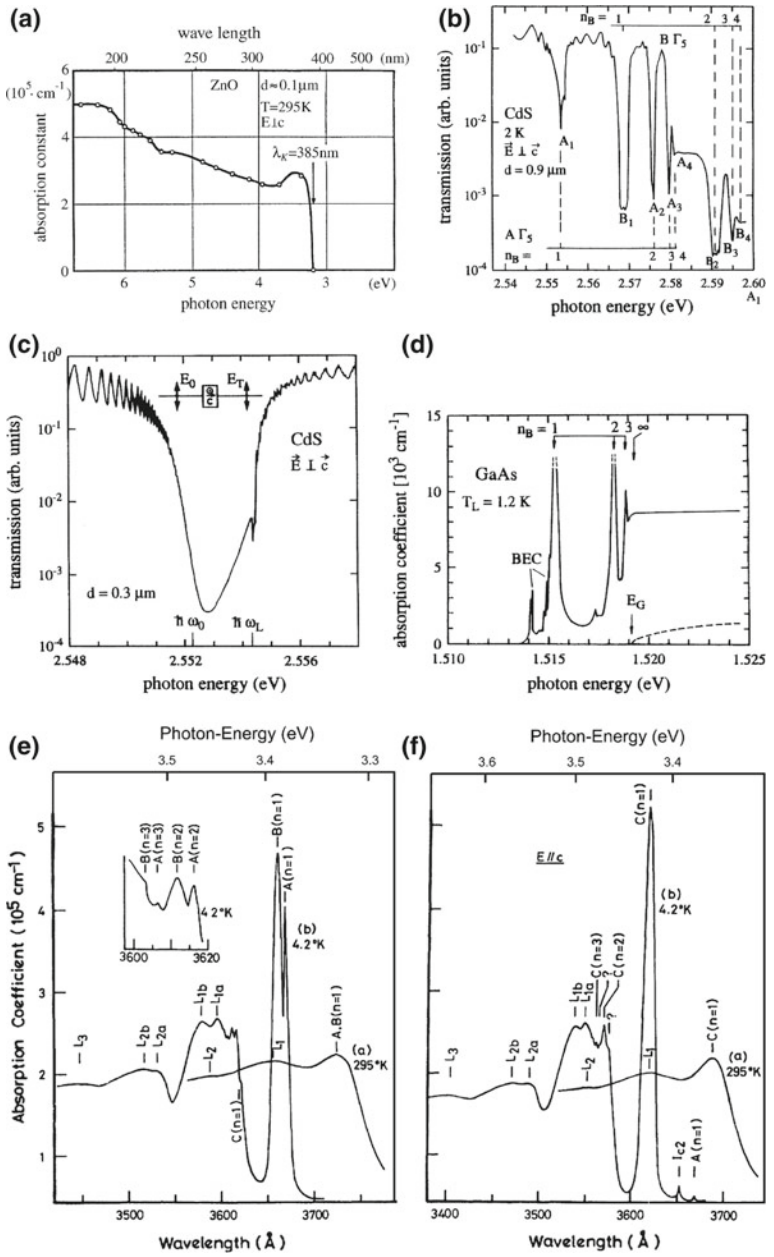


Fig. 21.8 An early experimental absorption spectrum of ZnO at RT (**a**), an experimental transmission spectrum of CdS for the orientation $\boldsymbol{E} \perp \boldsymbol{c}$ in the region of the A and B exciton resonances (**b**); a calculated one for the $A\Gamma_5$ resonance for $n_B = 1$ (**c**); an absorption spectrum for a thin GaAs sample (**d**) and absorption spectra of ZnO for both polarizations and for 4.2 K and 295 K (**e**,**f**) (adapted with permission from [59H1] ©1959 Elsevier, from [79V1] ©2006 John Wiley and Sons, from [82R1], from [91U1] ©2013 John Wiley and Sons, and from [68L1] ©1968 American Physical Society http://doi.org/10.1103/PhysRevLett.20.59, respectively)

At higher temperatures the finer details of the absorption spectra are washed out due to the increasing homogeneous broadening resulting from the scattering with thermally excited phonons. One is then left essentially with the broadened $n_B = 1$ exciton states followed by their ionization continuum. Compare Fig. 21.8a with Fig. 21.8e for 295 K.

The rather constant value of $\alpha(\omega)$ in the region of the continuum states comes from the product of the square-root joint density of states (the dashed line in Fig. 21.8d gives the calculated absorption spectrum for a simple band-to-band transition without Coulomb effects for comparison) and the Sommerfeld factor, already discussed in Sect. 20.2 for $d = 3$.

A useful formula for fit of absorption spectra which includes the correlation of the electron–hole pairs for both bound and continuum states in a suitable way is the *Elliott formula*. It has been derived by Elliott [57E1] analogous to the Hydrogen problem for transitions between non-degenerate bands in the effective-mass approximation. This theoretical approach treated direct (momentum-conserving) transitions in direct-gap semiconductors as well as indirect-gap semiconductors. The Elliott formula can also be derived from the semiconductor Bloch equations (see, e.g., [04G1]).

The Elliott formula has been applied to evaluate absorption spectra of III–V semiconductors like GaAs [62S1] or GaP [71S2]. A generalized version which can be extended to account for band non-parabolicity and broadening of the excitonic states has recently been introduced in [14S1] and applied to fit absorption spectra of hybrid organic/anorganic trihalide perovskite films:

$$\alpha(\hbar\omega) \propto \frac{\mu_{cv}^2}{\hbar\omega} \sum_{n_B} \left|\phi_{n_B}(r=0)\right|^2 \delta(\hbar\omega - E_{n_B}) \tag{21.10}$$

$$\propto \frac{\mu_{cv}^2}{\hbar\omega} \left[\sum_{n_B} \frac{4\pi\sqrt{E_b^3}}{n_B^3} \delta(\hbar\omega - E_{n_B}^b) + \frac{2\pi\sqrt{E_b}\,\Theta(\hbar\omega - E_g)}{1 - e^{-2\pi\sqrt{\frac{E_b}{\hbar\omega - E_g}}}} \right]$$

Here, μ_{cv}^2 is the dipole moment of the band-to-band transition, E_b and $E_{n_B}^b$ are the binding energies of the fundamental and and higher excitonic bound states, the factor $\left|\phi_{n_B}(r=0)\right|^2$ weights the density of states by the probability to find electron and hole at the same position. The terms in the second row of (21.10) describe the contribution of bound (delta-functions)and continuum states, respectively. Θ is the Heaviside step function.

An example for application of (21.10) is given in Fig. 21.9 for the novel solar-cell absorber material methylammonium lead iodide ($MAPbI_3$) [18R1]. The excitonic nature of the below-gap absorption peak is confirmed by the spectral coincidence of the modulus spectrum calculated from the electro-absorption spectrum (see 16.2). It is obvious again that the absorption is strongly deviating from the simple square-root law with a high absorption already at E_g. The absorption stays significantly enhanced due to the excitonic enhancement for photon energies far above the gap even at room temperature [18R1].

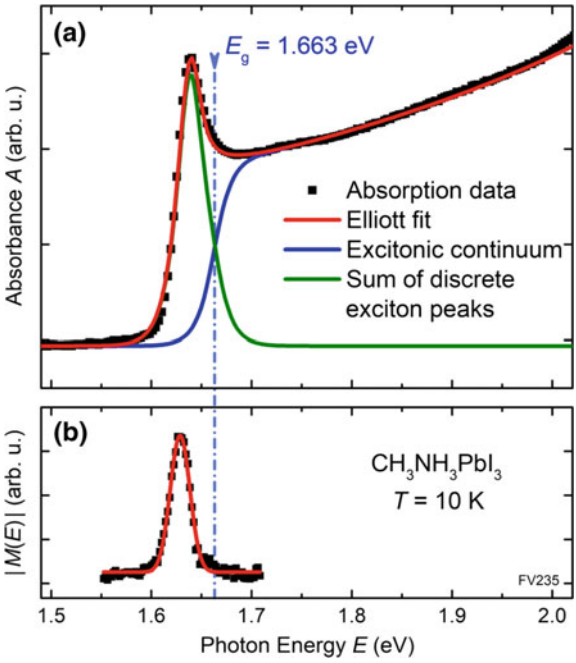


Fig. 21.9 Experimental absorption spectrum of methylammonium lead iodide and Elliott fit (**a**) using the extended version of (21.10) (see [14S1]) compared to the modulus spectrum |M(E)| from electro-absorption measurements (**b**) (reprinted from [18R1] with permission of AIP Publishing https://doi.org/10.1063/1.5017943)

At higher temperatures, the excitons develop an absorption tail to lower photon energies, which is described by the so-called *Urbach* or *Urbach–Martienssen rule* [53U1, 57M1, 58D1, 71K1, 71S1, 72D1, 73J1, 85L1]

$$\alpha(\hbar\omega) = \alpha_0 \exp\left[-\sigma(T)\,(E_0 - \hbar\omega)/\,k_B\,T\right] \quad \text{for} \quad \hbar\omega < E_0\,, \tag{21.11}$$

where α_0 and E_0 are material parameters. E_0 is an energy situated several 10 meV above the energy of the lowest free exciton at $T_L = 0\,K$, σ is a function varying only weakly with temperature. Examples are given in Fig. 21.10 for CdS or for ZnO in [73J1].

The temperature dependence of the bandgap energy, which can also be nicely seen in Figs. 21.6, 21.8 or 21.10, is already incorporated in (21.11) since the exciton peak shifts in parallel to the gap.

The reason for the behavior described by (21.11), which is of rather universal nature in semiconductors and insulators, is the interaction of excitons with optical phonons. Two effects are usually discussed in theory: a momentary localization of the excitons in the randomly fluctuating field of optical phonons, or an ionization in their electric field. These two effects seem to contribute with a weight that depends on the material parameters. Details about the theory can be found in [71S1, 72D1, 85L1].

There has been a certain misconception in parts of the literature concerning the interpretation of the low energy absorption tail in semiconductors, especially in the

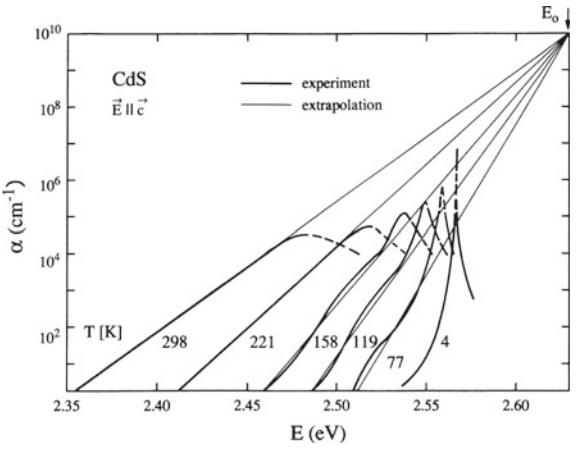


Fig. 21.10 The absorption edge of CdS for various temperatures and the polarization $\boldsymbol{E} \parallel \boldsymbol{c}$ (adapted with permission from [76S1] ©2006 Wiley and Sons)

work on ZnO and its alloys. This point has been addressed in detail in Chap. 6 of [10K2] and in [07K1, 10K1]. We summarize the main point here. We have understood already, that a simple square-root absorption edge starting at the bandgap is never observed in semiconductors due to the electron–hole Coulomb interaction. The latter results in exciton states below the gap and modifies the band-to-band transitions above the gap, i.e., in the excitonic ionization continuum (see Sect. 20.2). Furthermore, we have learned that an absorption tail develops below the exciton resonance due to interaction with phonons (the Urbach tail, see (21.11) above). Later in Sect. 22.2 we shall see, that disorder contributes additionally to this tail. Several authors tend now to fit this essentially exponential tail with either a simple square root absorption edge or a formula valid for the Tauc regime (see Sect. 18.2), extrapolate this fit to absorption coefficient zero and claim that this value gives the band gap or an "optical gap". It is clear from our knowledge of the near-edge absorption spectrum of semiconductors, that this approach is fundamentally wrong. So it is not surprising, that this fit coincides with the experimental data indeed usually only over a very limited range of absorption coefficients (frequently less than a factor of 3) and that the thus obtained values of the bandgap are systematically too small compared to the true value. The true values are frequently rather precisely known (e.g., for many III–V or II–VI compounds, see reference [82L1]). Furthermore the absorption coefficient measured at the thus obtained "bandgap" is orders of magnitude smaller compared to the correct value which is for ZnO e.g., around $2 \cdot 10^5$ cm^{-1} as seen from Fig. 21.8a, e, and f. We shall come back to these aspects shortly in Sect. 22.2.

The absorption related to direct excitons in indirect semiconductors will be addressed in Sect. 21.2.2.

21.1.3.3 Luminescence Spectra

The investigation of the luminescence from excitons or more precisely exciton polaritons is a rather difficult task. For early investigations see [72B2, 72W2]. The emission from these states is generally very weak even at low temperatures even in high quality sampels. This has various reasons: the total luminescence yield η_{lum} of semiconductors is often very low. Frequently one finds, even for direct gap materials $10^{-1} \gtrsim \eta_{\text{lum}} \gtrsim 10^{-3}$, i.e., the main recombination channel is non-radiative involving defect centers. A large part of the emission stems from phonon replica and, especially at low temperatures, from bound-exciton complexes, donor–acceptor pairs, or other defect centers which have been considered in Sect. 18.1. Furthermore, the direct emission from free exciton polaritons is limited by various effects: One is the internal reflection, another one the small escape depth. These points will be further clarified below.

If we excite an electron–hole pair, e.g., in the continuum states, it will relax to lower energies and thermalize by emission of phonons, as will be described in more detail in the second volume of this book. At very low temperatures ($k_{\text{B}}T < \Delta_{\text{LT}}$) the excitons end up on the LPB where they further relax by acoustic phonon emission. Since the scattering matrix element and the density of final states both decrease in the transition region between exciton- and photon-like dispersion, the excitons accumulate there. This is the reason why this region is called a bottle neck. At higher temperatures the excitons reach essentially a Boltzmann-like distribution on the exciton-like part of the LPB, on the longitudinal branch, and on the UPB. In the polariton picture the luminescence from these states cannot be described as a "recombination" of the exciton polariton with emission of a photon, since the photon will be immediately reabsorbed to form an exciton, or in other words, since we are considering the quanta of the mixed state of exciton and photon. Strictly speaking, an exciton polariton cannot simply recombine radiatively.

The proper description is the following: The exciton polariton moves with its group velocity through the sample. It can be scattered by impurities or phonons or be trapped. Eventually it reaches the surface. In most cases it will be reflected back into the sample. The limiting angle for total internal reflection α_{TR} is, e.g., for $n = 5$, which is a typical value on the LPB in the bottle-neck region as shown in Sect. 21.1.4, only about 13°. Of the excitons impinging under an angle smaller than α_{TR} a considerable fraction are also reflected back into the sample, as becomes clear if one considers the formulas for the reflection under normal incidence or if one integrates over Fresnel's formula of Sect. 5.1.4.

Furthermore, the luminescence yield of free-exciton polaritons is limited by the small escape depth l_{esc}, i.e., the depth from which they can reach the surface. If one excites in the band-to-band transition region, the exciting light penetrates about 0.1–1 μm into the sample corresponding to α-values of 10^4–$10^5\,\text{cm}^{-1}$ in this spectral range. By diffusion, the excitons spread out over a region of 1–2 μm. The depth from which they can reach the surface is much less than this. A rough estimate can be obtained either from the inverse effective absorption coefficient in the exciton resonance or from the product of phase-relaxation time T_2 (essentially the scattering

time) and the group velocity in the exciton resonance (Sect. 21.1.4):

$$\alpha_{\text{eff}}^{-1} = (10^4 - 10^6\,\text{cm}^{-1})^{-1} = 0.01 - 1\,\mu\text{m} \tag{21.12a}$$

$$l_{\text{esc}} = v_g T_2 = (10^{-3} - 10^{-5} \cdot c)\,(10 - 40\,\text{ps}) = (10 - 0.03)\,\mu\text{m} \tag{21.12b}$$

In spite of all these difficulties it was possible to observe the emission from the exciton polariton in many semiconductors and even its temperature dependence [84M3]. We give an example for ZnO in Fig. 21.11. On the left the dispersion relation of the A-exciton is shown for $\boldsymbol{k} \perp \boldsymbol{c}$ and $\boldsymbol{k} \parallel \boldsymbol{c}$ and for the polarizations $\boldsymbol{E} \perp \boldsymbol{c}$ and $\boldsymbol{E} \parallel \boldsymbol{c}$. Since the Γ_7 and Γ_9 valence bands are inverted in ZnO as compared to other Wurtzite type II–VI and III–V semiconductors like CdS, CdSe or GaN, the k-linear term appears in the A-exciton for the orientation $\boldsymbol{k} \perp \boldsymbol{c}$. The influences of the resulting additional polariton branch, of the longitudinal branch, and of the UPB are seen by comparison with the orientation $\boldsymbol{k} \parallel \boldsymbol{c}$ where the longitudinal branch and the k-linear term are missing. A luminescence polarized $\boldsymbol{E} \parallel \boldsymbol{c}$ in the A-exciton resonance observed in Fig. 21.11e, f is attributed to the AΓ_1 exciton, which is dipole allowed for $\boldsymbol{E} \parallel \boldsymbol{c}$ but has in ZnO a much smaller oscillator strength because it involves in this material a spin flip. Consequently both the spectral width and the longitudinal-transverse splitting are much narrow. See also [81K2, 10K1, 10K2] and references therein. A similar luminescence has been observed also in [03C1, 04K4]. Some other examples for the polariton luminescence including also other materials, both as bulk samples and epitaxial layers are found e.g., in [68V1, 74S1, 74T1, 75K1, 76S3, 77V1, 79P1, 80D1, 83M1, 95S1, 97N1, 97N2, 99R1, 00T1, 01R1, 01S1, 07K1, 10K1, 10K2]. The influence of a k-linear term has been discussed where applicable e.g., in [76S3, 82K1] or in the references given with Fig. 21.11.

A luminescence channel of the exciton polaritons with higher luminescence yield at low temperatures are the LO-phonon replicas [68S1] which show up in semiconductors with strong exciton–LO phonon coupling. In this case a polariton on the exciton-like part of the dispersion relation or in the bottleneck is scattered onto the photon-like branch by emission of one or more longitudinal optical phonons. The coupling with this type of phonon is stronger than with transverse optical or acoustic phonons since the lattice distortion of the polaron (Sect. 14.6) can be described largely as a superposition of longitudinal optical phonons. Once the polariton is on the photon-like branch, it travels over long distances with almost negligible damping and is transmitted through the surface into vacuum with quite high probability, since $n \lesssim 2$ in the corresponding spectral range.

In Fig. 21.12 we show the appearance of the LO-phonon satellites schematically. If we neglect the bottleneck region and homogenous broadening for the moment, we can deduce with the Boltzmann occupation probability the distribution of the excitons as a function of their kinetic energy E_{kin}

$$N(E_{\text{kin}}) \propto \begin{cases} E_{\text{kin}}^{1/2} \exp(-E_{\text{kin}}/k_{\text{B}}T) & \text{for } E_{\text{kin}} \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{21.13}$$

with $E_{\text{kin}} = \hbar^2 \boldsymbol{k}^2 / 2M$.

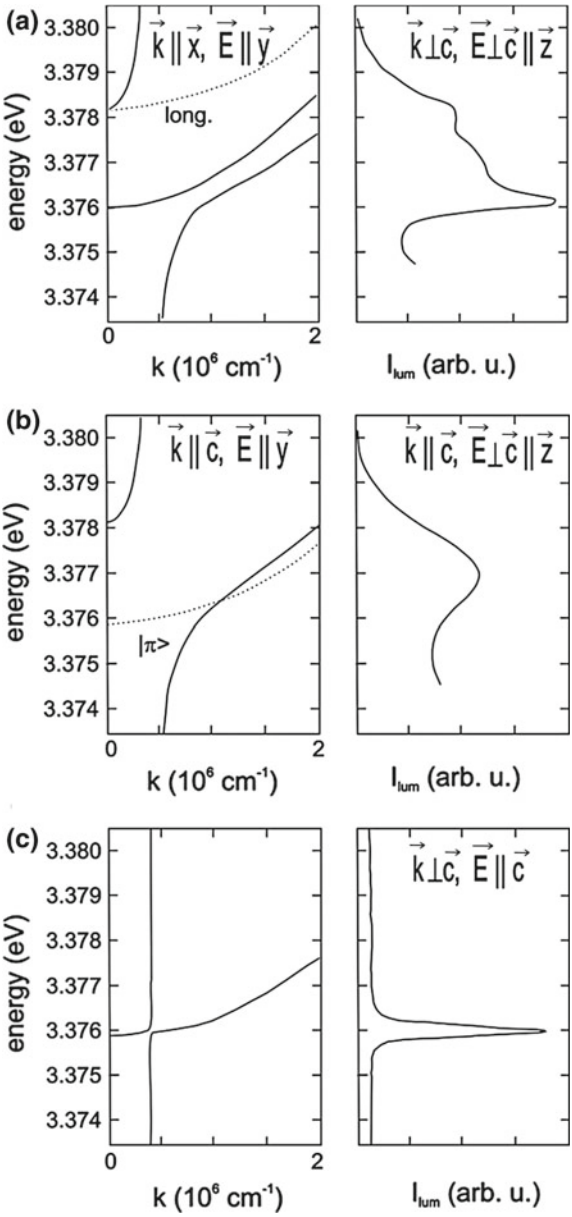


Fig. 21.11 The dispersion of the $n_B = 1$ A-exciton resonance in ZnO for various polarizations $\boldsymbol{E} \perp \boldsymbol{c}$ (**a**–**d**) and orientations $\boldsymbol{k} \perp \boldsymbol{c}$ (**a**, **b**) or $\boldsymbol{k} \parallel \boldsymbol{c}$ (**c**, **d**) and for $\boldsymbol{E} \parallel \boldsymbol{c}$ (**e**, **f**) (According to [78H1])

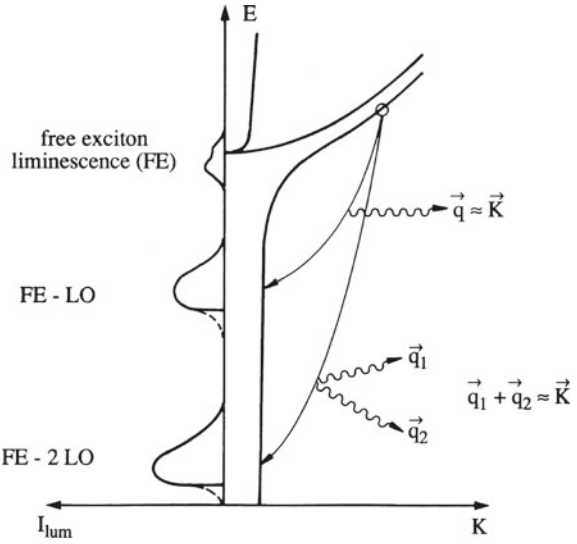


Fig. 21.12 Schematic drawing of the decay mechanisms of the exciton-mLO phonon emission processes

The lineshape of the luminescence of the mth LO-phonon replica is then given by [82P1, 06H1, 07K1]:

$$I_m^{\text{lum}}(\hbar\omega) \propto \begin{cases} E_{\text{kin}}^{1/2} \exp(-E_{\text{kin}}/k_{\text{B}}T)\, W_m(E_{\text{kin}}) & \text{for } E_{\text{kin}} \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{21.14}$$

with $\hbar\omega = E_0 - m\hbar\omega_{\text{LO}} + E_{\text{kin}}$, where E_0 is the energy of the dipole allowed, transverse exciton at $\boldsymbol{k}=0$. For the explicit discussion of side-band luminescence in uniaxial crystals see [97G1].

The *transition probability* $W_m(E_{\text{kin}})$ can often be expressed by a power law [82P1]:

$$W_m(E_{\text{kin}}) \propto E_{\text{kin}}^{l_m}\,. \tag{21.15}$$

For $m = 1$ one finds $l_1 = 1$ since the density of final states for the LO phonons increases with $E_{\text{kin}} \propto k^2$ assuming that the wave vector of the photon-like exciton polariton in the final state is negligible. For $m = 2$ many different combinations of the two-phonon wave vectors are possible for a given $\boldsymbol{k}$ of the exciton-like polariton. As a consequence l_2 is zero and the lineshape of the second LO phonon replica directly reflects the distribution of exciton polaritons in the initial state.

Finally I_m^{lum} has to be convoluted with a Lorentian, describing the temperature dependent homogenous broadening $\Gamma(T)$ as shown, e.g., in [06H1, 07K1, 07K2, 10K1, 10K2]. In Fig. 21.13 we show the luminescence of ZnO for various temperatures. At 110 K (Fig. 21.13a) the bound exciton emission around 3.35 eV is almost gone (compare to Fig. 22.2) and the near-edge luminescence is dominated by the homogeneously broadened zero phonon lines of the emission, which are still spectrally resolved, and by their LO phonon replica. The fit coincides nicely with

experiment. This coincidence confirms very well the concepts developed above and especially that the excitons are good quasi-particles at low excitation densities (the situation at medium and high excitation densities will be discussed in detail in the second volume). The distribution of the excitons on their dispersion relation can be described in many cases by Boltzmann statistics in close analogy to an ideal gas. A further example will be given for Cu_2O in Sect. 21.2. The value of the homogeneous broadening deduced from the Lorentzian fit of the zero-phonon emission has been used in Fig. 21.13a–c in the convolution of the LO phonon replica (21.13). The energy of the LO phonon (72 meV) is seen in the energetic distance from the zero-phonon line to the low-energy sides of the X-mLO replica, typically the points of inflection.

With further increasing temperature the homogeneous broadening increases (see Fig. 21.13b,c) and a discrepancy between experiment and theory develops on the high energy side which is due to the re-absorption within the temperature dependent Urbach tail. At room temperature the zero-phonon emission and its replica merge to an almost unstructured band with a maximum at (3.29 ± 0.02) eV and a FWHM of (90 ± 10) meV. Larger half widths reported by some authors at RT are indicative for additional inhomogeneous broadening presumably due to poor sample quality. Note that the maximum of the emission at RT is not identical with the free-exciton energy at this temperature.

The homogeneous broadening (HWHM, half width at half maximum) is plotted in Fig. 21.13d as a function of temperature and with the fit function [90R1] and parameters given in this figure. The value of the homogeneous broadening (which must not be confused with the longitudinal transverse splitting) of the A and B Γ_5 exciton polaritons starts at 0 K with a value Γ_0 around 1 meV for high quality bulk samples and epilayers but increases rapidly with temperature and reaches at RT a value of more than 20 meV HWHM! See for a detailed discussion of the parameters e.g., [07K1, 10K1, 10K2].

Finally, the temperature dependence of the bandgap of ZnO deduced from luminescence and absorption data is given in Fig. 21.13e. The fit follows the models described in [99P1]. As a rule of thumb one can state that the gap of ZnO decreases from 0 to 300 K by (90 ± 10) meV.

From (21.14) and (21.15) one can deduce that the ratio $Q_{1,2}$ of the integrated intensities of the first and the second LO-phonon replica is proportional to T

$$Q_{1,2} = \frac{\int I_1^{\mathrm{lum}}(\omega)\,\mathrm{d}\omega}{\int I_2^{\mathrm{lum}}(\omega)\,\mathrm{d}\omega} \propto T\,. \tag{21.16}$$

In Fig. 21.14 we give experimental data for $Q_{1,2}$ in ZnO for volume excitation and surface excitation. The first case has been realized by two-photon excitation with a ruby laser, which allows relatively homogeneous excitation of samples up to thicknesses in the mm range, and the second by UV excitation in the continuum states where the excitation depth is limited mainly by diffusion to values of the order of μm as discussed above. Up to temperatures of 100 K the points follow nicely the predictions of (21.16), then they drop. This deviation is due to the mentioned

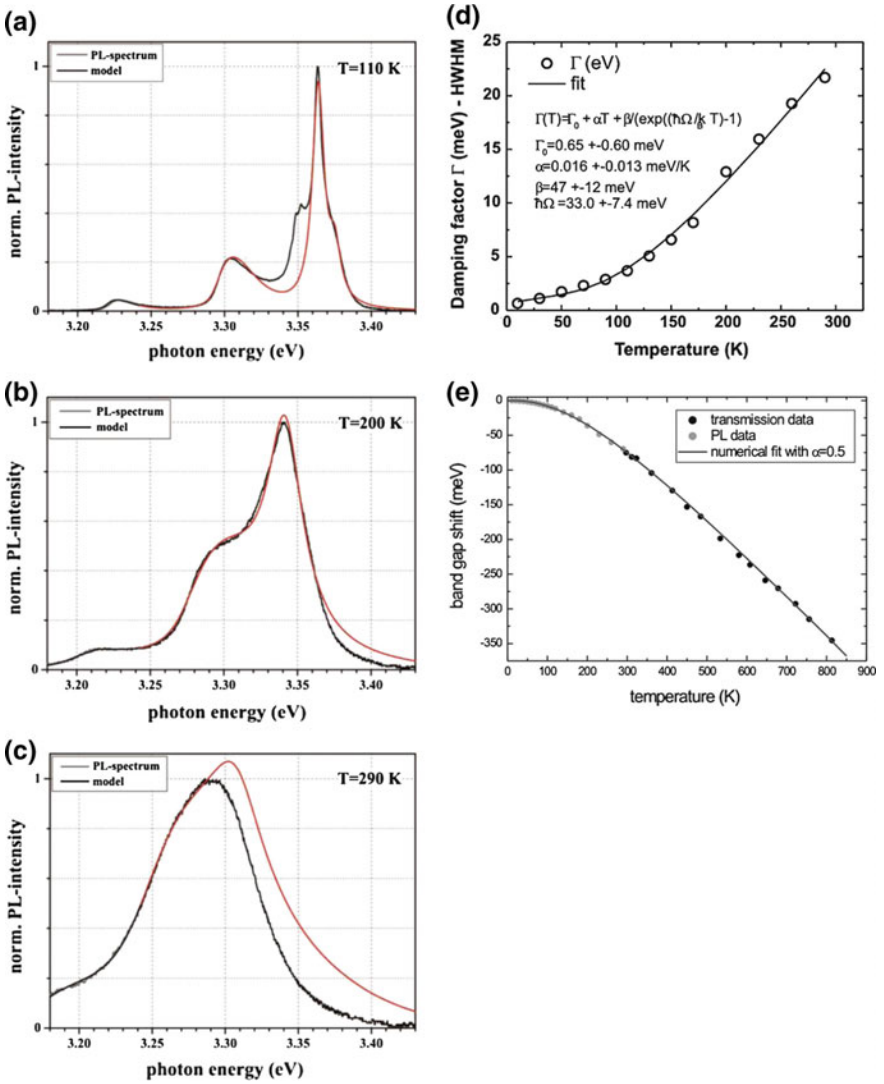


Fig. 21.13 The near edge luminescence of ZnO excited $c\omega$ in the band-to-band transition region for various temperatures with model calculations (**a**–**c**), the damping Γ of the A and B $n_B = 1\Gamma_5$ excitons as a function of temperature together with a fit function and fit parameters (**d**) and the temperature dependence of the gap with a fit function (**e**) (reprinted with permission from [10K1] ©2010 Wiley and Sons, based on data from [06H1, 07K1, 07K2])

re-absorption effects caused by the absorption tail described by (21.11), which starts to influence the escape depth of the polaritons also in the $m = 1$ range at higher temperatures. This effect is evidently more pronounced for volume excitation than for surface excitation. In semiconductors with less polar coupling, such as GaAs, the LO-phonon replicas are less pronounced.

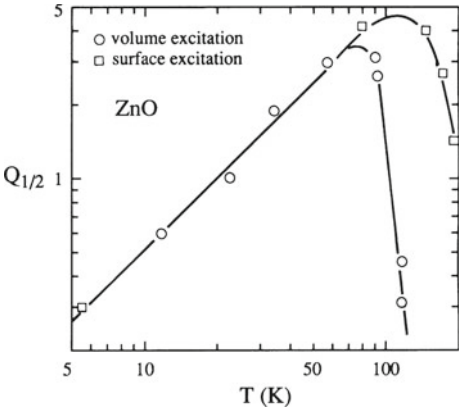


Fig. 21.14 The ratio of the integrated intensities of first and second LO-phonon replica in ZnO as a function of temperature for two different excitation conditions (according to [75K1])

A well-established concept used in the description of luminescence from ions in insulators is the *Huang–Rhys factor* S. It should be mentioned here that for free excitons in most semiconductors S is below one even for ones with strong ionic binding. The large ratio of first to zero-phonon intensity frequently observed is not connected with a large value of S but with other processes that reduce the zero-phonon emission, as outlined above (see e.g., [68V1, 79P1]). In [03Z1] the Huang-Rhys factor has been deduced from the fit of the first and second LO-phonon replica in ZnSe even as a function of the kinetic energy of the exciton-like polaritons. It decreases from values around 0.3 for small $\boldsymbol{k}$ vectors to 0.1 for a kinetic energy of about 20 meV.

LO-phonon replicas do not only appear in the excitonic luminescence spectra on the low energy side, but may also appear in case of sufficiently strong exciton LO-phonon coupling in the absorption spectra on the high energy side, extending into the continuum. Examples have been cited above in connection with the term exciton-phonon bound state [66R1, 68T1, 72B1, 72K1, 94W1].

It is obvious from Figs. 21.11 or 21.13 and from the way in which the line shape of the X-mLO replica has to be fitted that a deconvolution into Gaussians is inadequate. A homogeneously broadened luminescence band has usually a Lorentian line shape, an inhomogeneously broadened one may have incidentally a Gaussian one, but there are no reasons that it has to be Gaussian, and a comparable influence of homogeneous and inhomogeneous broadening may lead to a Voigt profile. Nevertheless, one finds frequently a deconvolution of luminescence bands or even of nonlinear response functions (see e.g., references cited in [07K1, 10K1, 10K2]) into Gaussians. The reason is possibly the easy availability of the corresponding subroutine, but not arguments of physics.

The interaction of excitons and photons with acoustic phonons in Cu_2O even leads to a three-component eigenstate called *phonoriton* [99H1, 00H1].

It should be mentioned that excitons can also be scattered by emission of quasi-particles other than phonons onto the photon-like part of the polariton branch. An example is, e.g., scattering by magnons, which may occur in (anti-)ferromagnetic semiconductors like $RbMnF_3$ or Cr_2O_3 [76S2, 99D1].

To conclude this discussion of free-exciton luminescence, it should be mentioned, that calculations [98K1] show, a luminescence feature may occur at the position of the free-exciton resonance, resulting, however, from the recombination of an electron–hole plasma (see volume two) going under emission of a photon from a state containing n electron–hole pairs to one with $(n-1)$ pairs. Though these calculations do by no means rule out the existence of excitons as good quasi-particles, a luminescence feature at the exciton energy alone is not necessarily a proof of their existence. The observation of additional features, like intra-excitonic transitions (see Sect. 21.3 and [03K1, 06K1]) can help to clarify possible discrepancies. In these two references many semiconductors with small exciton binding energy have been investigated like GaAs, which show with increasing density and/or temperature readily ionization and transition to an electron–hole plasma. In materials with larger exciton binding energy like the Cu-halides, the wide gap II–VI semiconductors, the group III nitrides or Cu_2O excitons will better survive as individual quasi-particles, though also here ionization occurs at sufficiently high densities and/or temperatures (see [07K1, 07K2, 10K1] and references therein).

As a rule of thumb it can be stated that the formation of excitons is likely for resonant and low or intermediate exciton densities. If the thermal energy $k_{\mathbf{B}}T$ is comparable to or larger than the exciton binding energy $E_{\mathrm{x}}^{\mathrm{b}}$ the excitons will be rapidly thermally ionized and have a large homogeneous width. See Fig. 21.13e. Under band-to-band excitation, exciton formation is likely for $k_{\mathrm{b}}T \ll E_{\mathrm{x}}^{\mathbf{b}}$ and for exciton formation times shorter than the electron–hole pair lifetime. The formation time is short, if the excess energy can be dissipated by LO-phonon emission especially in polar materials, while acoustic-phonon emission is slower, especially in less polar materials. We will discuss these point in more detail in volume two of this book.

Experimental Technique: Integrating Sphere and Luminescence Yield

In order to achieve quantitative values for, e.g, the quantum efficiency of radiative recombination (absolute *luminescence yield* = number of emitted photons/number of absorbed photons), but also for separation of absorption of light and reflective losses at the sample surface one would need an absolute calibration of the complete experimental set-up. It is obvious that it is nearly impossible to account in a reproducible way for spectral through-put or the efficiency of all optical components like detector, monochromator/spectrometer, lenses, mirrors, beam-splitters, etc., and for the geometry like the solid angle of acceptance of the spectrometer, the size of the image of the excitation spot on the entrance slit of the spectrometer and the spatial radiation characteristics of the sample.

One could compare a measured luminescence spectrum of the sample under investigation with a standard, which has luminescence efficiency close to unity (usually a laser dye or some luminescence centers in a solid matrix). This necessitates identical geometries for both experiments including the excitation

depth and the spatial radiation characteristics. These conditions are usually hard to fulfill.

The way out of these problems is placing the sample under investigation in an *integrating* or *Ulbricht sphere*. This geometry and a high-reflective coating of the sphere's interior eliminate the radiation characteristics of the sample and lead to a homogeneous illumination of the interior [98W1]. This allows to obtain even absolute, external luminescence efficiency per spectral interval with rather simple calibration techniques. Fitted into a cryostat one can also determine activation energies and thus get information on excited states and/or non-radiative recombination channels [08H1].

For low-temperature integrating spheres one uses the thermoplastic resin spectralon which has a diffuse reflectance of >99% for the UV-VIS-NIR spectral region. Three ports are located on the equator of the sphere (see Fig. IS-1). The baffle, also made of spectralon, protects the detection port from the direct irradiance from the sample. Either cw light sources or pulsed lasers can be used for optical excitation. The light homogenized inside the sphere consists of the photoluminescence as well as the reflected and scattered pump light from the sample surface and the sample holder. A tiny fraction is coupled out of the sphere through a small hole (detection port) and spectrally analyzed.

By illuminating the sphere wall with a known photon flux, the detector sensitivity is determined, that means the sphere wall is used as a reflection standard. Since the fraction of light coupled out of the sphere into the detection system is within a small variance independent of the wavelength and the emission characteristics of the sample (see [98W1]), one has direct access to the absolute external quantum efficiency.

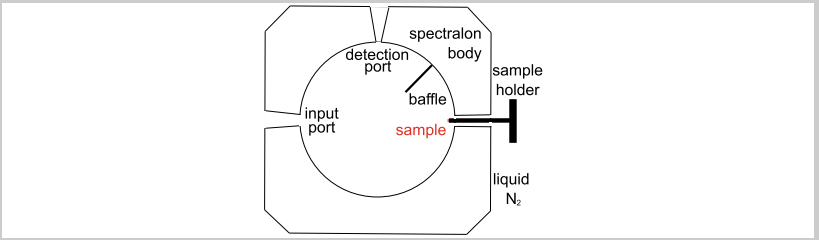


Fig. IS-1 Cross-section (top view) of a low-temperature integrating sphere made of spectralon. The sphere is fitted into a cryostat and cooled down to 77 K via thermal contact with a LN_2 reservoir. The sample holder is in thermal contact with a liquid He reservoir allowing sample temperature down to 6 K (for details see [98W1])

Examples for the usage of integrating spheres are found in [90L1, 93V1, 97W2, 98W1, 03F1, 08H1]. An experimental example is shown in Fig. IS-2 for a ZnSe/ZnMgSSe multiple quantum well sample. The figure shows the external quantum efficiency $\eta_{\mathrm{lum}}(T)$ of the spectrally integrated luminescence from the quantum well as well as the one from the barrier. The dependence on lattice temperature T is fit by a thermal activation model:

$$\eta_{\mathrm{lum}}(T) = \left(1 + \frac{w_{\mathrm{nr}}}{w_{\mathrm{r}}} + \sum_{i=1}^{N} \frac{w_{\mathrm{nr}i}^{a}}{w_{\mathrm{r}}} \cdot \exp\frac{-E_{\mathrm{a}i}}{k_{\mathrm{B}}T}\right)^{-1}, \qquad \text{(IS-1)}$$

where w_{r} and w_{nr} are the radiative and nonradiative recombination rates, respectively, and E_{ai} are activation energies. The latter relate to delocalization of excitons and to quenching due to exciton ionization by optical phonons. For the well luminescence an additional activation energy is given by the thermal activation of excitons into the barrier [98W1].

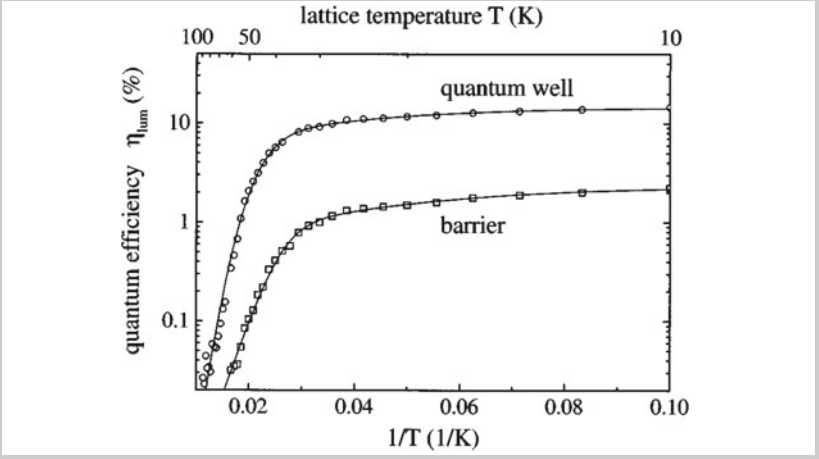


Fig. IS-2 Temperature dependent quantum efficiency of luminescence from wells and barriers of a ZnSe/ZnMgSSe multiple quantum well. The solid lines are fits using a thermal activation model (reprinted from [98W1], with the permission of AIP Publishing)

An alternative to the integrating-sphere method is to measure the heat deposited in the sample under investigation. If the absorbed energy is known, it is easy to calculate the luminescence efficiency. Examples for this *calorimetric absorption spectroscopy (CAS)* are found in [81K3, 88J1, 91B1, 94F2]. However there are some draw-backs: the limitation very low temperatures, the requirement to have sufficiently sensitive bolometers, and the fact that the luminescence may be emitted in a spectral range which is not detected, e.g., via deep centers resulting in an over estimation of the luminescence yield of the monitored luminescence channels. For further thermal methods see e.g., [81J1, 82I1].

Another general caveat is that often *differential efficiency* is given in literature (partly without explicitly mentioning it), e.g., the slope in the relation between emitted number of photons versus excited electron–hole pairs. If this relation is strictly linear through the origin, the absolute efficiency and the differential efficiency are the same. If the relation is nonlinear, e.g., above the threshold of stimulated emission of a laser diode, the differential efficiency may be close to unity while the absolute efficiency is still much smaller.

21.1.4 Spectroscopy in Momentum Space

It is clear from the above discussion that the concept of exciton polaritons allows one to understand the spectra of reflection, transmission and luminescence, but that a quantitative interpretation of the data usually involves a rather elaborate theoretical fit. Therefore various techniques have been developed which allow more or less directly the spectroscopy of exciton polaritons in momentum space, i.e., they provide the possibility of measuring the dispersion relation $E(\boldsymbol{k})$ more directly.

We start this section by recalling briefly the consequences of *spatial dispersion*, i.e., of the dependence of ω_0 on $\boldsymbol{k}$, and then we present various methods of $\boldsymbol{k}$- or *momentum-space spectroscopy*.

As discussed in Chap. 8, the combination of the dielectric function $\varepsilon(\omega, \boldsymbol{k})$ (8.14) and the polariton (2.36) forms an implicit representation of the polariton dispersion. For $\omega > \omega_{0,\mathrm{L}}$ there are two propagating modes in the sample or even more if the dispersion relation is complex (Figs. 8.9, 21.3 or 21.11) or if the longitudinal branch couples to the radiation field as may occur for $\boldsymbol{k} \neq 0$ or for oblique incidence in uniaxial crystals. Since the $\boldsymbol{k}$-vectors of the various modes in the sample are different, the diffracted beams propagate in different directions, thus giving some meaning to the term "spatial dispersion". Below $\omega_{0,\mathrm{L}}$ we have at least one propagating and one evanescent mode, which however, for finite damping also acquires a small real part. The consequences of this fact for the reflection spectra have already been discussed above.

The first method of $\boldsymbol{k}$-space spectroscopy we want to illustrate uses the analysis of the *Fabry–Perot modes* introduced in Sect. 5.3. Such modal structure appears for example in as-grown, thin, platelet-type samples with plane-parallel surface. For an example see Fig. 21.8b. As already pointed out there, transmission maxima occur when an integer number of half-waves fit into the resonator, i.e.,

$$k_m(\omega_m) = m\frac{\pi}{d}\,,\quad m = 1, 2, 3\ldots\,, \tag{21.17}$$

where d is the geometrical thickness of the sample and k is the real part of the wave vector in the medium. If ω_m and k_m are known for one particular m then the dispersion relation can be reconstructed from (21.17) by reading the $\hbar\omega_m$ from the transmission spectrum and by progressing in steps of πd^{-1} on the k-axis. In Fig. 21.15 we show the dispersion of the $n_{\mathrm{B}} = 1$ AΓ_5-exciton polariton in CdSe, together with a measured and a calculated reflection spectrum. The condition for the reflection minima coincides with that for the transmission maxima in (21.17).

The equidistant spacing on the k-axis is clearly visible and the good agreement of theory and experiment is obvious. Above $\omega_{0,\mathrm{L}}$ one can clearly see closely spaced modes of the LPB and superimposed widely spaced ones from the part of the light traveling on the UPB. The fact that the small modulation decreases with increasing photon energy indicates that the fraction of light traveling on the LPB through the sample decreases with increasing energy above $\omega_{0,\mathrm{L}}$.

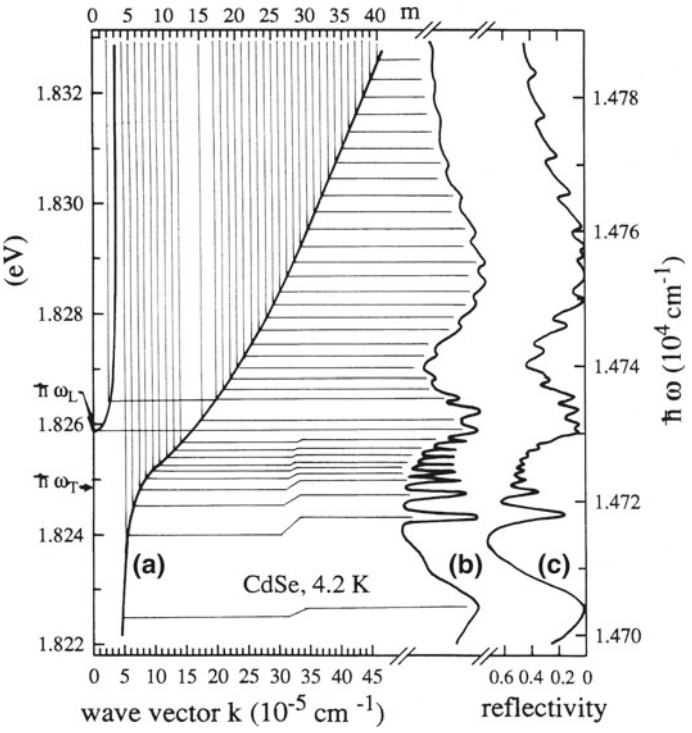


Fig. 21.15 The polariton dispersion in CdSe with an equidistant ruling of the $\boldsymbol{k}$-axis (**a**) and the observed (**b**) and calculated structure (**c**) of the resulting Fabry–Perot modes (adapted by permission from [75K2] ©2006 Wiley and Sons)

Similar data have also been found in CdS and CuCl [80M1, 84M1, 84M2, 93K1]. For polariton propagation in shallow confinement heterostructures see e.g., [96N1, 97N2, 04S1] and references therein.

The next method is *resonant Brillouin scattering*. We introduced this scattering process with acoustic phonons in Sect. 12.2 and recall that the Brillouin shift is directly proportional to the transfer of momentum due to the linear dispersion relation of acoustic phonons around the Γ point.

Figure 21.16a shows schematically the Brillouin scattering in the resonance region of an exciton polariton in a backward scattering configuration. The effect has been observed in GaAs, CdS, ZnSe and CuBr [77W1, 78U1, 79S1, 80B3, 80K1, 81C1, 84M1, 84S1, 84W1]. In Fig. 21.16b, c we give data for CdS and GaAs partly involving TA and LA phonons. The soild lines have been calculated in a fit procedure with the parameters in the inset. These of course, also allow the calculation of the polariton dispersion itself.

A close inspection of the Brillouin line shape shows that the damping constant is energy or $\boldsymbol{k}$-dependent [84S1]. This result has also been obtained in CuCl from non-

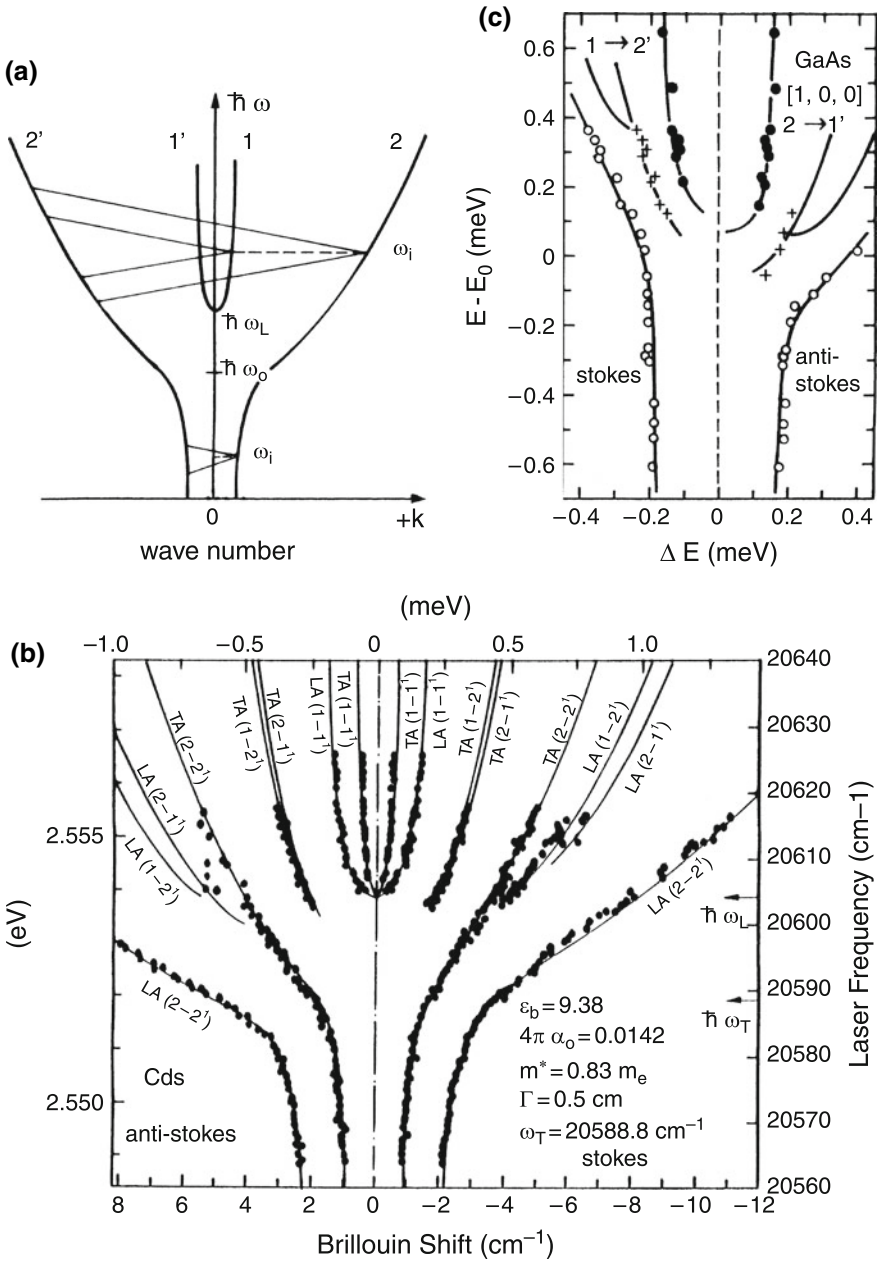


Fig. 21.16 Schematic drawing of Brillouin scattering in backward geometry in the resonance of an exciton polariton (**a**) and experimental data for GaAs (reprinted with permission from [78U1] ©1978 Springer Nature) (**c**) and for CdS (reprinted with permission from [84W1] ©1984 American Physical Society http://doi.org/10.1103/PhysRevB.29.3350) (**b**)

degenerate four-wave mixing [85H1]. We will describe this technique in detail in the second volume. Similarly one finds that Δ_{LT} is also a function of $\boldsymbol{k}$ [82O1, 85H1].

It is well known that refraction from a prism can be used to determine the real part of $\tilde{n}$ and thus of $\boldsymbol{k}$. If sufficiently thin ($d \lesssim 1\,\mu$m), prism-shaped samples are available, it is also possible to extend this technique to the resonance region of the exciton polariton. Fortunately some CdS platelets grow in the desired form, presumably involving a small-angle grain boundary.

Figure 21.17 gives two examples for CdS, where n is given and not k, but where both quantities are simply connected with each other according to $\mathrm{Re}\{k\} =$

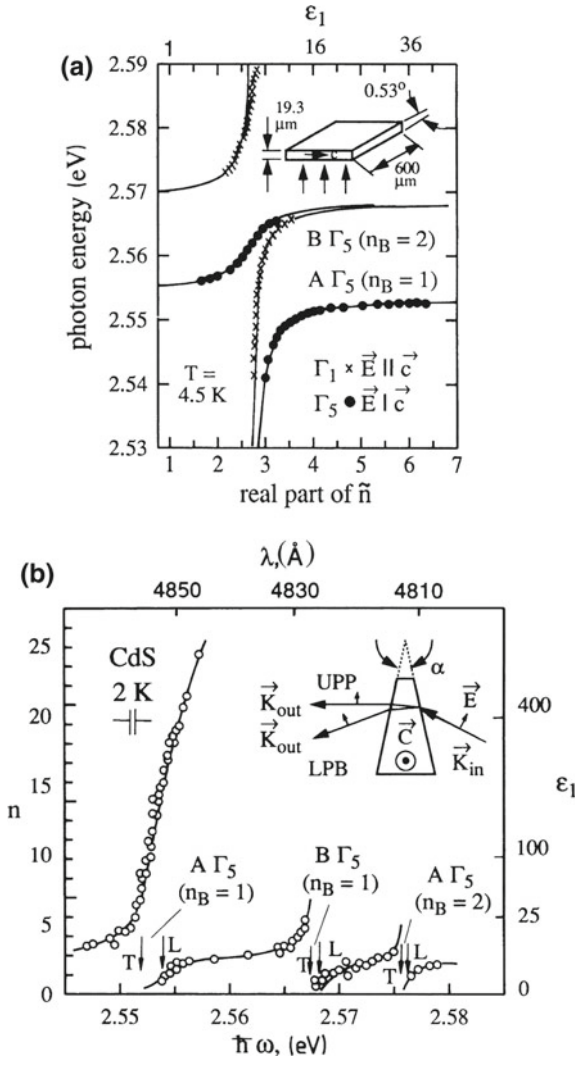


Fig. 21.17 Dispersion relations of the A- and B-exciton resonances determined with the *thin-prism method* (reprinted with permission from [81B1] ©1981 Elsevier and [84L1])

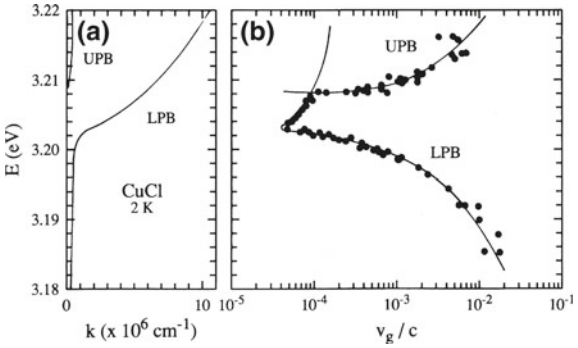


Fig. 21.18 The dispersion relation of the lowest free exciton resonance in CuCl (**a**) and the group velocity determined from the time-of-flight of picosecond laser pulses and from the dispersion relation (**b**) (reprinted with permission from [79M1] ©1979 The Physical Society of Japan)

$\omega/c\,\mathrm{Re}\{\tilde{n}\}$, see (2.39). In Fig. 21.17a one sees again the birefringence and dichroism for the polarizations $\boldsymbol{E} \perp \boldsymbol{c}$ (i.e., Γ_5) and $\boldsymbol{E} \parallel \boldsymbol{c}$ (i.e., Γ_1) of the $n_\mathrm{B} = 1$ A-exciton resonance known already from the reflection spectra of Fig. 21.3. In Fig. 21.17b the dispersion of the LPB can be followed up to $n \approx 20$ corresponding to $\varepsilon_1 = 400$ or $k = 2.5 \cdot 10^6\,\mathrm{cm}^{-1}$.

If one compares the *time of flight* of a picosecond pulse through a sample with its propagation through vacuum one can deduce the group velocity v_g and, with (2.45), the slope of the dispersion relation. In Fig. 21.18b experimental values are given for v_g at the lowest exciton polariton resonance in CuCl, together with a curve deduced from the dispersion relation of Fig. 21.18a. The excellent agreement between experiment and theory again proves the validity of the polariton concept. One can see from Fig. 21.18b that in CuCl v_g can be as low as $5 \times 10^{-5}\,\mathrm{c}$.

In the region above $\omega_{0,\mathrm{L}}$ two pulses are created in the medium from one incident pulse due to the different group velocities on the LPB and the UPB. The spatial distance between them increases with sample thickness and this is another reason to call the $\boldsymbol{k}$-dependence of ω_0 "spatial dispersion".

Similar experiments have been performed in GaAs, CdSe and CdS [79U1, 81I1, 83S1, 01N1]. In [01N1] the phase shift has been observed in GaAs, if one passes over the exciton resonance. It is essentially the analog of the phase shift between excitation and resonator in classical mechanics in a driven or forced oscillation if one passed with the excitation frequency from below (phase shift zero) through the resonance (phase shift $\pi/4$) to values above it (phase shift $\pi/2$).

We conclude this subsection with two nonlinear methods of $\boldsymbol{k}$-space spectroscopy, namely *two-photon spectroscopy* or *absorption (TPA)* and *two-photon* (or *hyper-*) *Raman scattering* (*TPRS, HRS*), anticipating some of the results discussed in the second volume of this text book.

Experimental Techniques: Two-Photon Absorption and Hyper Raman Scattering

In two-photon absorption (TPA) simultaneous absorption of two photons with energies $\hbar\omega_{1,2}$ and momenta $\hbar\boldsymbol{k}_{1,2}$ occurs. These two photons can stem from two laser beams which are directed onto the sample, sometimes with a finite angle between them. But also absorption of two photons with identical energies $\hbar\omega$ and momenta $\hbar\boldsymbol{k}$ from a single sufficiently strong laser is possible. Necessary condition for the TPA process is that the sum of energies and momenta coincide in the sample with those of an excited state, i.e., (see Fig. TPA-1a)

$$\begin{aligned} \hbar\omega_1 + \hbar\omega_2 &= E_\mathrm{f}\,, \\ \hbar\boldsymbol{k}_1 + \hbar\boldsymbol{k}_2 &= \hbar\boldsymbol{k}_\mathrm{f}\,. \end{aligned} \qquad \text{(TPA-1)}$$

Due to the curvature of the LPB it is only possible with TPA to reach states on the longitudinal exciton branch or on the UPB (see Fig. TPA-1). The selection rules are usually different from those of one-photon absorption and depend in addition on the polarizations of the photons relative to each other and to the crystallographic axes. Examples will be given in volume two. Even three-photon spectroscopy has been used to determine the exciton-polariton branches in various semiconductors and insulators [85B1, 86B1, 93F1, 94F1]. See also Fig. 21.20.

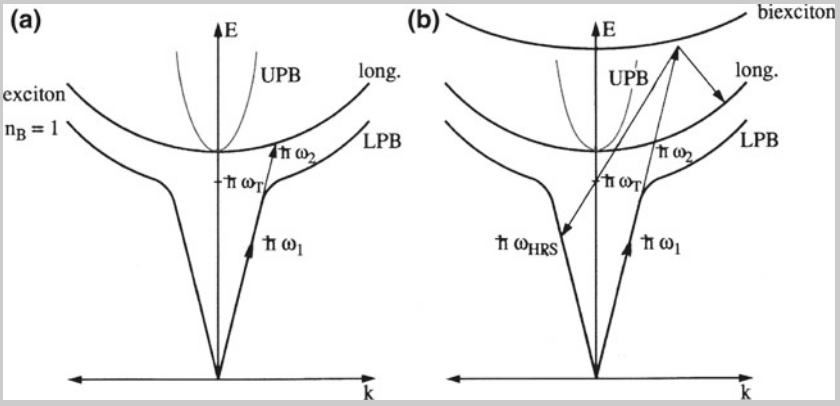


Fig. TPA-1 Two-photon transition from the crystal ground state to the longitudinal exciton branch (**a**) and the hyper-Raman process which is almost resonant with the exciton and the biexciton state (**b**)

The hyper-Raman scattering (HRS) process starts from two (generally photon-like) polaritons which can couple coherently to a *two-polariton state*. This virtual state decays under energy and momentum conservation into a photon-like polariton $\hbar\omega_\mathrm{R}$, which is observed as a Raman-like emission, and another

final-state particle $\hbar\omega_f$ which can be photon- or exciton-like. If a single incident beam is used to deliver both incident quanta, i.e., $\hbar\omega_1 = \hbar\omega_2$, this process reads

$$2\hbar\omega_1 = \hbar\omega_R + \hbar\omega_f \qquad \text{(TPA-2)}$$
$$2\hbar\boldsymbol{k}_1 = \hbar\boldsymbol{k}_R + \hbar\boldsymbol{k}_f\,.$$

It is advantageous to have one or both of the incident quanta almost in resonance with the exciton and/or the biexciton state (Sect. 20.4). This choice enhances the transition probability due to the small resonance denominators appearing in perturbation theory.

In Fig. TPA-1 we show schematically both a TPA and a HRS process where the longitudinal branch acts as final state. Corresponding experimental data are presented in Fig. TPA-2. Here the polariton dispersion in CuCl around the lowest $n_B = 1$ Γ_5-resonance and the states reached by TPA and by HRS are shown.

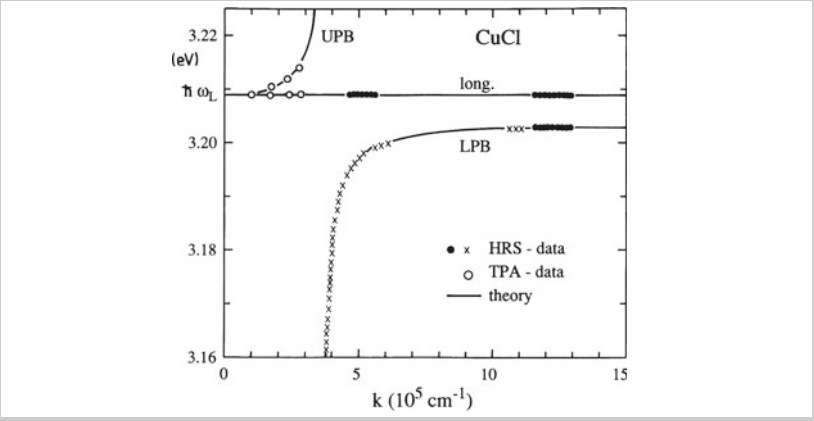


Fig. TPA-2 The polariton dispersion of the lowest Γ_5-exciton resonance in CuCl measured by two-photon absorption and by hyper-Raman scattering (reprinted with permission from [81K1] ©1981 Elsevier)

A large number of semiconductors have been investigated by TPA and HRS. TPA has also been used to detect higher exciton states ($n_B > 1$). For reviews see, e.g., [71F1, 81K1, 85H1, 97W1]. Three-photon absorption has proved the validity of the polariton concept also in insulators like KCl [85B1, 94F1, 97W1] as mentioned above.

A slightly more involved observation of the polariton dispersion with short and thus spectrally broad pulses ($\Delta_{LT} < \Delta_{Pulse}$) has been already addressed in Sect. 8.3 in the context of *propagation quantum beats* between lower and upper branch polaritons. The example of Cu_2O, where the optical transition across the direct gap is dipole-forbidden, can be complemented by data for the dipole-allowed case like ZnSe or InSe [96N1, 97N1].

Momentum-space spectroscopy is also an important tool in the topical field of quantum electrodynamics of cavities (*cavity-QED*). Here the polariton concept has been revived to describe the strong coupling between photonic modes of an optical resonator and matter excitations (like excitons). These *cavity polaritons*, [06K2, 08L1, 10W1, 11G1] which even show macroscopic coherent phenomena like polariton condensation and superfluidity, will be discussed extensively in volume two of this text book.

21.1.5 Surface-Exciton Polaritons

As already discussed in Sect. 8.6, one finds surface-exciton polaritons in the spectral region between the LPB and the longitudinal exciton. Their dispersion can be measured by *attenuated* (or *frustrated*) *total internal reflection* (*ATR*).

Experimental Technique: Attenuated Total Reflectance

Internal reflection and attenuated total reflectance (ATR) are mainly known in the context of ATR infrared spectroscopy which is a powerful tool to investigate, e.g., strongly absorbing liquid or soft-matter samples with applications in chemistry, bio-medicine or forensic science. For reviews see [67H1, 93M1, 12M1]. Here we will describe the application of ATR to excite and investigate surface polaritons.

Since surface polaritons cannot propagate into medium I (generally vacuum) they cannot be created by shining light of an appropriate frequency on the sample. The same is true for the other side. As a consequence, it is not possible to excite surface polaritons directly. The method of ATR involves attenuated total reflection at a succession of two the interfaces. A light beam is sent into a prism in such a way that total internal reflection occurs at its base (Fig. ATR-1a). Under these conditions an evanescent wave leaks into the vacuum as is indicated schematically (see also Fig. 8.13). The frequency of this wave can be tuned trivally, and its wave vector $\boldsymbol{k}_{\parallel}$ by varying the angle of incidence. If a semiconductor is brought close enough to the base of the prism, the evanescent wave can couple to the surface-exciton polariton modes in the semiconductor if both $\hbar\omega$ and $\boldsymbol{k}_{\parallel}$ coincide. "Close enough" means distances of the order of λ, i.e., fractions of a μm and the realization of this condition involves some experimental skill. The coupling to the surface-exciton polariton mode attenuates the total reflection.

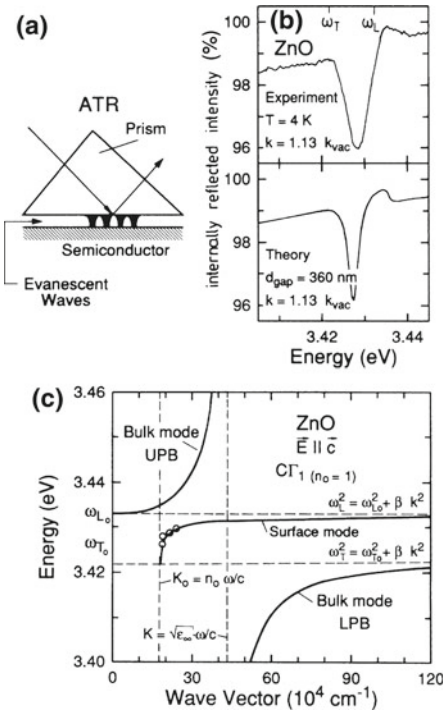


Fig. ATR-1 The principle of the experimental technique of attenuated total reflection (ATR) (**a**) experimental and calculated ATR spectra for the $n_B = 1$ $C\Gamma_1$-exciton in ZnO (**b**) and the resulting dispersion of the surface and volume polaritons for $n_0 = 1$ (**c**) (Adapted with permission from [78L1] ©1978 Springer Nature)

In Fig. ATR-1b measured and calculated ATR spectra are shown for the $n_B = 1C\Gamma_1$-exciton resonance in ZnO, which has rather large Δ_{LT} splitting of about 12 meV. The calculated dispersion relation of the surface-exciton polariton is given in Fig. ATR-1c together with the states reached by ATR. Good agreement between experiment and theory can be claimed in both Fig. ATR-1b, c.

More details about surface-exciton polaritons in various semiconductors including their observation in nonlinear optical spectroscopy can be found in [78L1, 80F1, 81L1, 82H1, 82S2, 84F1] or in [82M1].

It should be noted, that some authors use the term "surface exciton" for an exciton state bound to a defect at the surface or close to it (see e.g., [05G1]). These states occur energetically below the transverse eigenenergy of the free exciton, are a type of bound-exciton complex (see Sect. 22.1) and have nothing to do with surface-exciton polaritons. Some care is necessary not to confuse both notations.

21.1.6 Excitons in Organic Semiconductors, in Organic/Inorganic Hybrids and in Insulators

Organic semiconductors are nowadays found in various applications like displays, light-emitting diodes (OLEDs) or solar cells. Their advantages lie, e.g., in their high mechanic flexibility, in low-cost, often solution based processing techniques and the ability to print devices. Frequently organic semiconductors are polymeric materials made from carbon-based, chain-like, quasi one-dimensional molecules (see e.g., [02L1]). They feature conjugated binding in the carbon back-bone and weak (van der Waals) inter-molecular binding. But also crystalline materials like the model substance anthracene exist. Recently, organic/inorganic hybrid materials like the perovskite-type lead halides have drawn extensive attention as promising absorber materials for thin-film solar cells.

Crystalline organic semiconductors usually contain several molecules per unit cell, anthracene, e.g., four molecules of $C_{14}H_{10}$. Correspondingly complex is the spectrum of lattice vibrations. Sometimes it is possible to distinguish between *phonons* or intermolecular modes, i.e., modes where the molecules in the unit cell oscillate relative to each other in the sense of acoustic and optical modes and *vibrons* or intramolecular modes, where the atoms oscillate within one molecule. Obviously there may be some coupling between these two types of modes.

Concerning the electrons, there are valence bands called *highest occupied molecular orbitals (HOMO)* and conduction bands called *lowest unoccupied molecular orbitals (LUMO)* with a gap between them. The bands are usually rather narrow, due to the small electronic overlap between molecules, resulting in large effective masses of the charge carriers (electrons and holes) and low mobility ranging from 1 to 10 cm^2Vs.

The inter-band excitations in the electronic system form excitons, however with some characteristic differences from the (Wannier) excitons in most inorganic semiconductors. Usually one finds in organic semiconductors and especially in the wide gap ones *Frenkel excitons*, i.e., electrons and holes reside in the same molecule. The excitonic binding energy of oligomers (small molecules of identical units called monomers) and polymers (large chains of monomers) is mainly determined by the charging (or Coulomb) energy. It thus shows a universal behavior on the chain/conjugation length and increases strongly for small molecules [03K2].

The close spatial confinement of excitons results in singlet-triplet splittings which can easily exceed 1 eV (see, e.g., [82L1]). The triplet states have low oscillator strength, low luminescence yield and long lifetime in the ms range. The singlet states have high oscillator strength and a short life time in the ns-range and a tendency to relax into the lower lying triplet states (inter-system crossing). Depending on the nature of the chemical bonding, coupling of excitons to (Raman- and/or IR-active) phonons may be strong, leading possibly to *self-trapping* of the exciton [94H1]. The interation with the phonons leads in any case to a strong Stokes shift between absorption and emission spectra, which have approximate mirror symmetric appearance du to the Franck–Condon effect (see, e.g., [10F1]).

Excitons with large oscillator strength form polaritons in sufficiently pure materials in a similar way as in inorganic semiconductors.

Furthermore there are *charge transfer (CT) excitons*, in which the two carriers (i.e., electron and hole) reside in adjacent molecules forming, e.g., a pair of a positively and a negatively charged anthracene molecules.

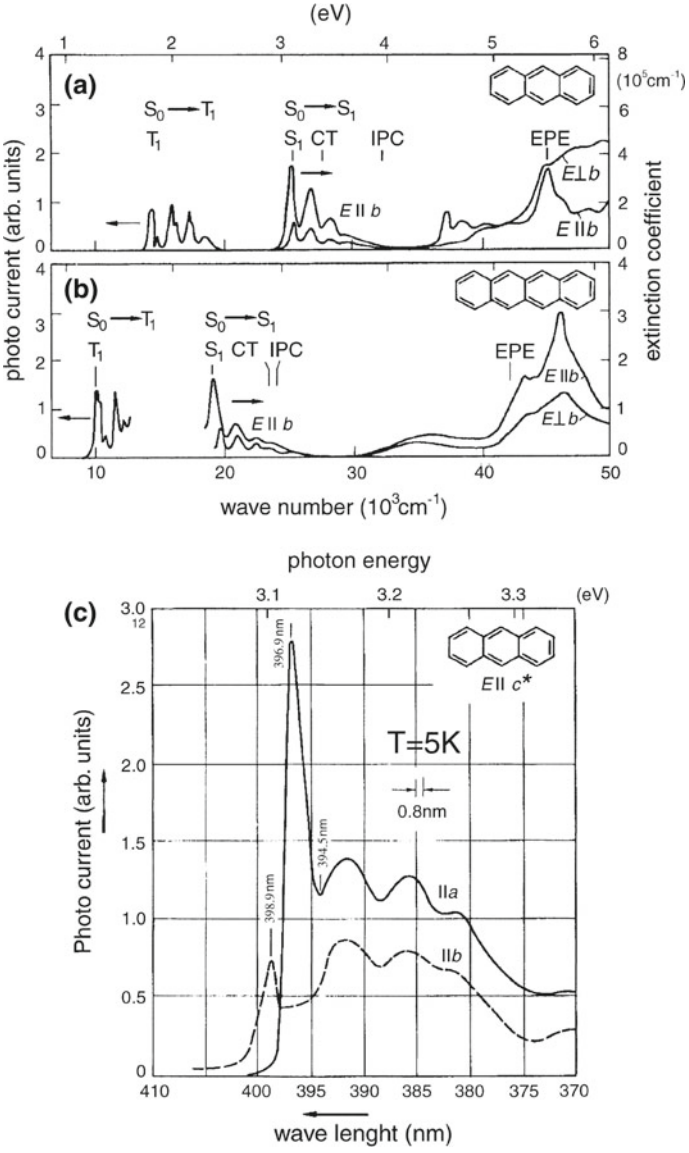


Fig. 21.19 Overview of the absorption and photocurrent excitation spectra of anthracene (**a**) and tetracene (according to [79S1])(**b**) and photocurrent excitation spectra (PCE) of the singulet exciton in anthracene (**c**) (reprinted with permission from [73S1] ©1973 Springer Nature)

We only want to give here two examples of the optical properties of excitons in organic semiconductors. For further publications of optical properties of organic semiconductors see e.g., [12B1, 09K1, 10T1].

In Fig. 21.19a,b we give an overview of the absorption spectra of anthracene and tetracene. One can see the transitions from the crystal ground states (S_0) to triplet (T_1) and singlet (S_1) excitons. The oscillator strength of the first transition is too low to observe it directly in absorption.

Therefore, the photoluminescence excitation spectra (see 23.2) of the delayed triplet fluorescence are plotted. The singlet states have, as already mentioned, a high oscillator strength resulting in values of the absorption coefficients exceeding $3 \cdot 10^5 \mathrm{cm}^{-1}$. CT stands for charge transfer exciton, IPC for the intrinsic photoconductivity threshold and EPE for the external photoemission threshold. The crystals are anisotropic, therefore the spectra are given for the polarizations parallel and perpendicular to the crystallographic b axis.

Figure 21.19c shows a close-up of the absorption spectrum of the singlet exciton polariton absorption, measured here with the technique of photo-current excitation spectroscopy. The latter works similar to the photoluminescence excitation spectroscopy (23.2) only that a photo-current rather than PL is detected.

Typical organic light-emitting materials used in displays and OLEDs are small molecules like Aluminumoxinate (Alq_3) [87T1], polymers like poly(p-phenylene vinylene) (PPV) and derivates thereof [90B1], copolymers of polyflourene (PFO) [14S2], poly-p-phenylene (PPP) [14S2] or dendrimers doped, e.g., with Ir complexes like $Ir(mppy)_3$ [04Y1]. For a review see [14S2].

Recently organic/inorganic hybrids like perovskite-type metal halides of composition $ABX_{3-x}Y_x$ (e.g., $A = CH_3NH_3^+$; $B = Pb^{2+}$; X,Y = I^-, Br^-, Cl^-) have been introduced as promising solar-cell absorbers or light emitting materials [14G1]. The exciton binding energies are are here much smaller than in pure organic semiconductors [15M1, 18R1]. An example of excitonic features in optical spectra of $CH_3NH_3PbI_3$ has been given in Fig. 21.9. Investigation of excitonic transitions in magnetic fields also reveal that the exciton binding energies are below the thermal energy at room temperature. [16G1, 17Y1]

These materials combine high optical absorption and long carrier lifetimes [17H1, 17M1] with bandgap energies [14C1] that can be optimized for highly efficient solar cells or for covering the full visible spectral range in light-emitting devices (see, e.g., [16S1, 16V1] for reviews).

As an example of the exciton polariton dispersion in an inorganic insulator, namely KI, we show the dispersion relation obtained from three-photon sum and difference frequency spectroscopy in Fig. 21.20. Note the high value of $\Delta_{LT} \approx 100$ meV. Further information including similar compounds, is found, e.g., in [85B1]. More information on excitons in insulators, including self trapping or solid rare gases can be found, e.g., in [64K1, 67K1, 67O1, 78G1, 83G1].

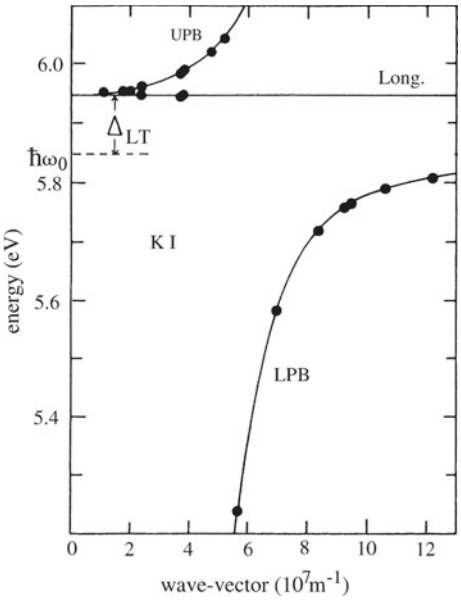


Fig. 21.20 The exciton polariton dispersion in KI obtained from three-photon spectroscopy including the lower and upper polariton branches and the longitudinal exciton (adapted with permission from [94F1] ©1994 Springer Nature)

21.1.7 Saddle-Point and Core Excitons

So far, this chapter has focused on electronic transitions in semiconductors close to the fundamental gap. But transitions from deeper valence bands and/or into higher conduction bands are also possible. The structures connected with these transitions are obviously found at $\hbar\omega > E_g$, i.e., in the (V)UV part of the spectrum for semiconductors with bandgaps in the visible.

The absorption in the range of the exciton continuum is influenced by the Sommerfeld enhancement as discussed in [91O1] and shown in Figs. 21.8d and 21.9. Its importance, i.e., the dependence of the Sommerfeld factor on the transition energy decreases with increasing energy above the onset of the excitonic ionization continuum (the bandgap) and the optical transitions are more or less determined by the band-to-band transitions.

We have discussed this scenario already in Sect. 16.2 where we introduced the concept of critical points in the band structure and their consequences for the density of states. For band-to-band transitions the spectral shape of the optical functions, i.e., of the reflectivity $R(\omega)$, of the complex dielectric function $\varepsilon(\omega) = \varepsilon_1(\omega) + \varepsilon_2(\omega)$ or of the complex index of refraction $\tilde{n}(\omega) = n(\omega) + i\kappa(\omega)$, is determined by the joint density of states (see (16.9)). Looking again at the spectra from Ge in Fig. 16.3 or GaAs in Figs. 16.2 and 16.2 we notice that neither the transition at the indirect gap of Ge nor the lowest M_0 critical point, which results in the distinct exciton lines shown for GaAs in Fig. 21.8c, contribute significantly to the overall spectra. The fact that the features in these spectra look generally more like peaks and less like the critical

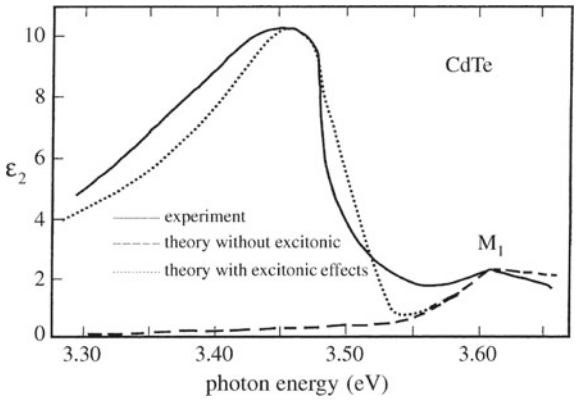


Fig. 21.21 The imaginary part of ε_1 in the region of the E_1 feature in CdTe compared with two calculations taking electron–hole Coulomb correlations into account or not (reprinted with permission from [10Y1] ©2010 Springer Nature, original data from [69K1, 70W1, 71P1])

point structures (e.g., saddle points) shown in Fig. 16.2 is to a large extent due to the neglect of the electron–hole Coulomb interaction in the latter.

The excitonic features resulting from the Coulomb correlation also modify the optical properties of the higher energy transitions in the form of *saddle-point excitons* and *core excitons*. However, these excitons are often not resolved as individual structures, even at low temperatures, since the damping of these states is too high (or the phase-relaxation time too short) due to the fact that they can decay rapidly into lower energy states.

To illustrate this statement, we give in Fig. 21.21 an example for CdTe. The direct gap is at 1.4 eV, the E_1 feature is located around 3.5 eV. The dashed line shows the expected spectrum for ε_2 bearing in mind the M_1 critical point in the joint density of states. The measured spectrum deviates significantly, but coincides nicely with a model calculation taking, in addition to the M_1 critical point, the Coulomb interaction between electron and hole, i.e., excitonic effects into account.

At much higher photon energies, namely in X-ray regime, the K, L, … absorption edges of the elements forming the semiconductor [93M2, 96T1] dominate the spectra. These deep core levels are influenced only slightly by the chemical binding, but the absorption spectrum is still influenced by the position and properties of the conduction band since the electron has to be excited into an empty state, or in other words, by the interference of excited electrons scattered from neighboring atoms. These facts are exploited in *extended X-ray absorption fine structure (EXAFS)* and *X-ray absorption near edge structure(XANES)* experiments. The absorption coefficient may still reach considerable values in the X-ray regime, but has an overall tendency to decrease with increasing photon energy as shown in Fig. 21.22. The imaginary part of $\tilde{n}$ is in contrast rather small in this region and the real part is very close to unity, approaching this value above the highest resonance in the material from below as predicted in Chap. 7.

In Fig. 21.22 we show these trends for NaCl. This is an insulator but this does not matter to much in the present context. The data compiled in Fig. 21.22 are remarkable, because they cover eight orders of magnitude in photon energy. They show nicely the optical-phonon resonance in the IR, the exciton and band edge in the UV slightly

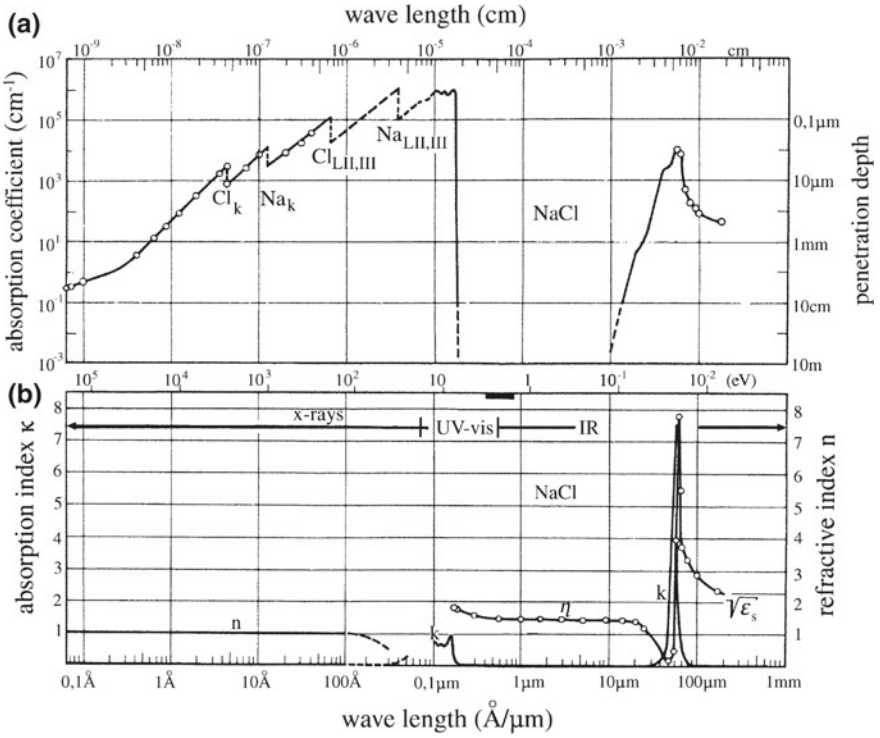


Fig. 21.22 The spectra of the absorption coefficient of NaCl (**a**) and of the real and imaginary parts of $\tilde{n}$ (**b**) (adapted with permission from [76P1] ©1976 Springer Nature)

below 10 eV and lead from there towards higher energies to the element specific absorption edges in the X-ray regime.

More information about core excitons and related topics can be found in [65C1, 65C2, 66G1, 85I1, 90K1, 93M2, 94J1, 95D2, 96T1, 99A1].

21.2 Forbidden Exciton Transitions

In the preceding section, we treated the optical properties of exciton resonances with large oscillator strength, requiring the polariton concept for a quantitative description. There are also intrinsic excitons which, for various reasons, couple only weakly to the radiation field. In this case the weak coupling approach is usually, but not always (see, e.g., Fig. 8.5c), sufficient for their description.

Reasons why excitons can have low oscillator strength include spin flips, dipole-forbidden transitions, or the necessity to involve a third particle, e.g., a phonon in their creation. We give some examples in the following.

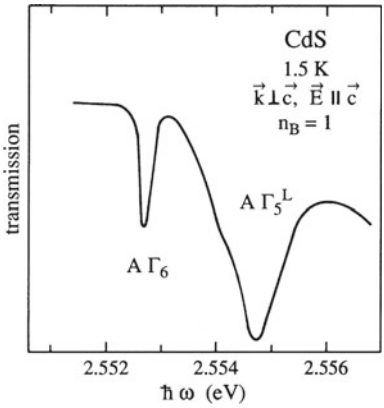


Fig. 21.23 A transmission spectrum of CdS in the resonance region of the $n_B = 1$ A-exciton for the orientation $\boldsymbol{k} \perp \boldsymbol{c}$ showing the dipole-forbidden transitions $A\Gamma_6$ and $A\Gamma_5^L$ (according to [82B2])

21.2.1 Direct-Gap Semiconductors

21.2.1.1 Triplet States and Related Transitions

Even in direct-gap semiconductors with dipole-allowed band-to-band transitions there are exciton states with very small oscillator strength, so-called forbidden transitions. Among them there are spin triplet states (see Sect. 20.2) which involve a spin-flip, longitudinal excitons, and states whose symmetry forbids the transition in the dipole approximation (Appendix B).

The latter situation can occur e.g., for $n_B \geq 2$ and L or $L_z = 1, 2 \ldots$. Such exciton states can sometimes be observed in one-photon transitions in higher-order perturbation theory, e.g., as electric quadrupole or magnetic dipole transitions. This, however, is only possible if no strong one-photon transition occurs in the same spectral range, which masks the weak transition. An example is again CdS. We have already stated that the $n_B = 1$ A-excitons are dipole allowed only for the polarization $\boldsymbol{E} \perp \boldsymbol{c}$, i.e., Γ_5. For $\boldsymbol{E} \parallel \boldsymbol{c}$ one can weakly see the $A\Gamma_6$ triplet or the $A\Gamma_5^L$. In the latter case the finite angle of aperture of every real light beam plays a role because small deviations from $\boldsymbol{k} \perp \boldsymbol{c}$ lead to a small oscillator strength of mixed-mode states according to (21.9). In Fig. 21.23 we show a transmission spectrum of CdS around the n_B=1 A-exciton states for the orientation $\boldsymbol{E} \parallel \boldsymbol{c}$ and $\boldsymbol{k} \perp \boldsymbol{c}$, where the $A\Gamma_6$ triplet and the $A\Gamma_5^L$ states are seen. A further example for ZnO will be given in Sect. 24.1.

An alternate way of reaching some of these states is via two-photon absorption (TPA). The selection rules for TPA differ from those for one-photon absorption as is outlined in Appendix B. A first example has already been presented in Fig. 21.20, where the longitudinal branch of the Γ_5 exciton is seen in TPA. Another example for states with $n_B \geq 2$ will be given in Sect. 24.1, where we discuss the influence of magnetic fields on excitons.

21.2.1.2 Parity Forbidden Band-to-Band Transitions

There is a group of semiconductors that have a direct gap at the Γ point. But the transition between the uppermost valence band and the lowest conduction band is dipole forbidden because both bands have the same parity. The best investigated material of this group is Cu_2O, but there are several others such as SnO_2, TiO_2, and GeO_2. In these materials exciton states with an S envelope (i.e., $L=0$) are all dipole forbidden. For $n_B \geq 2$ there are also envelope functions with $L=1$, which have odd parity. Via their envelope, these states acquire a weak oscillator strength and can be seen in absorption.

A model system for such materials and for exciton physics in general is Cu_2O. It is an early example of exciton research in semiconductors [52H1, 66K1] and also for reports on a possible Bose–Einstein condensation of excitons [93L1, 96M1, 97S1] (see volume two of this textbook). There has been recent work on its forbidden transitions [03D1] and on intra- excitonic transitions (see Sect. 21.3). Current focus is on the observation of excitonic states with high n_B, so-called *giant Rydberg excitons* [14K1].

Figure 21.24 gives in (a) the band structure of Cu_2O with the symmetries and parities of the conduction and valence bands and in (b) an example of the *yellow exciton series* in Cu_2O, where the one- and two-photon absorption spectra are compared. The yellow exciton series results from the transitions between the upmost valence- (Γ_7^+) and the lowest conduction band (Γ_6^+). Transitions involving deeper valence- and/or higher conduction bands lead to the green, blue and violet exciton series. The $n_B=1$ states of the yellow and green exciton series have been found by two-photon absorption. Their positions are indicated. These states do not fit into the n_B^{-2} series, for the reasons discussed already in Sect. 20.3. However, even for the tiny quadrupole oscillator strength of this $n_B=1$ S-exciton, it has been shown that the polariton effect exists and can be measured (see the polariton propagation quantum beats discussed in Sect. 8.3 [91F1]).

The 1S *ortho-exciton* of the yellow series, to which we concentrate in the following, has a binding energy of 140 meV while the nP ($n > 1$) states follow a hydrogen series with $Ry^* = 93.26$ meV (see, e.g., [98J1] and references therein). Due to the small exciton Bohr radius in Cu_2O of about 0.53 nm the singlet-triplet (or ortho-para) splitting is much larger than in usual semiconductors and amounts to 12 meV. The lower lying *para-exciton* is optically forbidden to all orders of perturbation theory. Its position has been deduced by applying magnetic (or strain) fields, which mix ortho and para states, and by extrapolating to zero field [79M1]. Additionally, it has been claimed that a weak Γ_5^- phonon replica of the para-exciton is seen in luminescence spectra around 2.01 eV [78B1].

In Fig. 21.25 we give some further examples of the optical properties of the yellow exciton series. The 1S ortho-exciton states around $\boldsymbol{k}=0$ show up only weakly in absorption or luminescence as seen in Figs. 21.24a and 21.25a, c, respectively. But the LO-phonon replica reproduce nicely the square-root density of states of this effective-mass particle. The onset of this Γ_{12}^- LO-phonon absorption band is shown together with a calculated square-root function in Fig. 21.25a. This is one of the few cases, where this square-root DOS of excitons (and not of the band-to-band

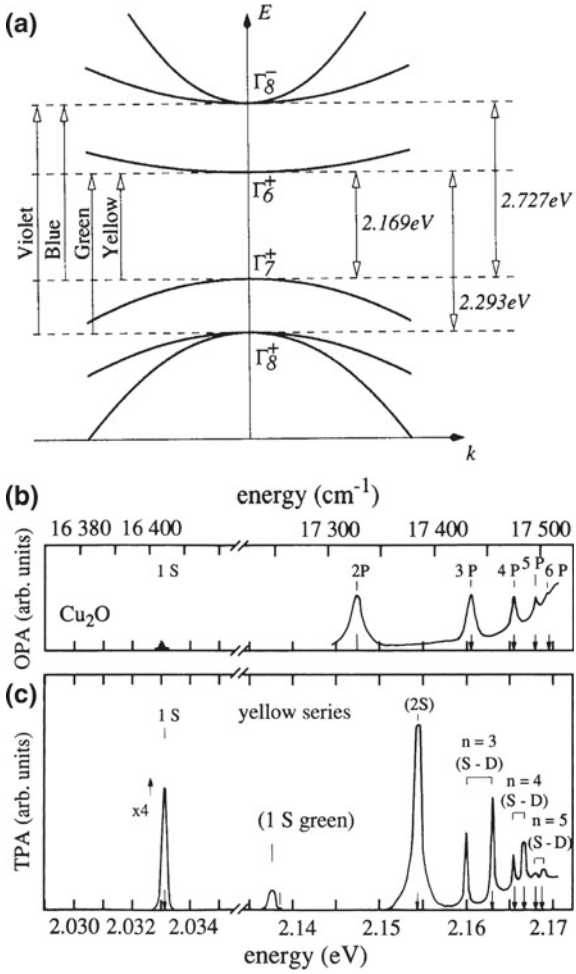


Fig. 21.24 The band structure of Cu_2O (**a**) and the absorption of the forbidden "yellow" exciton series in CuO_2 in one and two photon absorption (OPA and TPA, respectively) (**b**, **c**) (adapted with permission from [81U1] ©1981 American Physical Society http://doi.org/10.1103/PhysRevB.23.2731)

transition!) shows up directly in experiments in support of the concept of excitons as free quasi-particles.

Superimposed on this absorption band (also including higher bands) is the absorption of the *n*P ortho-excitons as shown in Fig. 21.25b. Their line shape is asymmetric resulting from the constructive and destructive interference of the transition amplitudes on both sides of the resonance, respectively, with the LO-phonon continuum in the sense of a Fano-type resonance.

The exciton luminescence is similarly dominated by the phonon replicas as seen in Fig. 21.25c showing both Stokes and anti-Stokes emissions. The long exciton lifetime results in thermalization of the excitons with the lattice. The theoretical curves in this figure are calculated essentially as a product of the square root DOS and a Bose–Einstein or Boltzmann occupation probability. The exciton temperatures

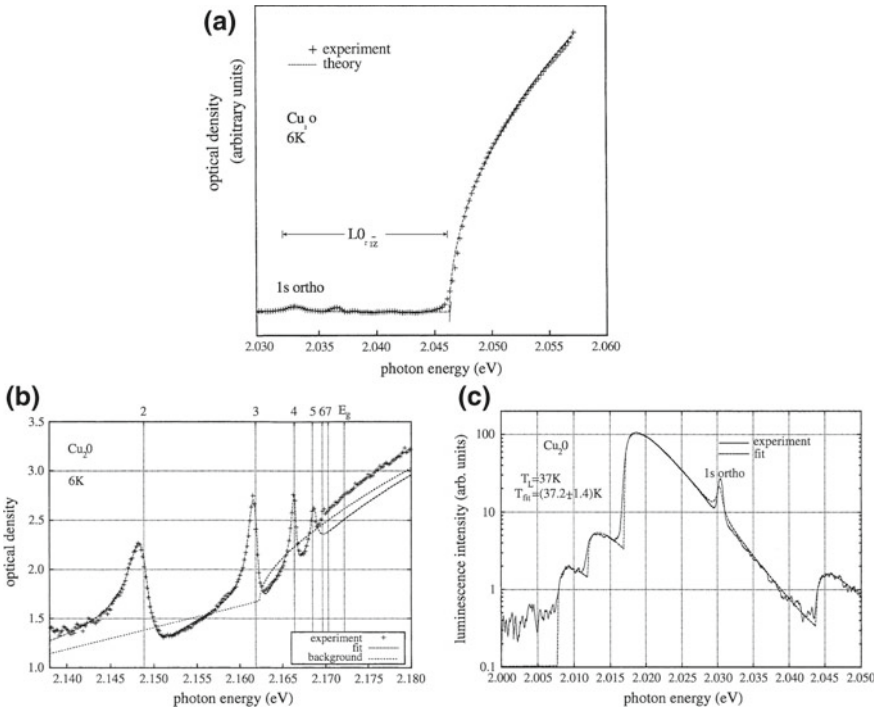


Fig. 21.25 Absorption (**a**) (according to [97J1]), **b**) and luminescence spectra (**c**) of Cu_2O (adapted with permission from [98J1] ©1998 John Wiley and Sons)

deduced from this fit and from the ratio of Stokes and anti-Stokes emission agree very well with the lattice temperature. This also holds for temperatures down to or even below 5 K [98J1].

The high exciton binding energy of Cu_2O facilitates the observation of *Giant Rydberg excitons* which are analogs of states in highly excited hydrogen atoms. Excitonic resonances with principal quantum numbers as large as $n_B = 25$ are found in natural crystals which have outstanding crystal quality by high resolution laser spectroscopy (see Fig. 21.26) [14K1]. Negligible carrier scattering at low excitation density and suppressed radiative transitions to lower exciton states lead to narrow resonances with an anisotropic lineshape due to the Fano effect described above. The diameter of these giant excitons reach values larger than 2 μm. The strong dipole–dipole interaction between the excitons results in a spatial blockade preventing the resonant excitation of further excitons within the extend of a present Rydberg exciton.

The Rydberg excitons show some similarities to hydrogen-like atoms but also subtle differences due to the occurrence in a solid-state environment. This environment allows observation of many highly excited states at readily available photon energies but also reduces the symmetry with respect to full rotational symmetry in the hydrogen atom. One finds effects due to the valence-band characteristics like

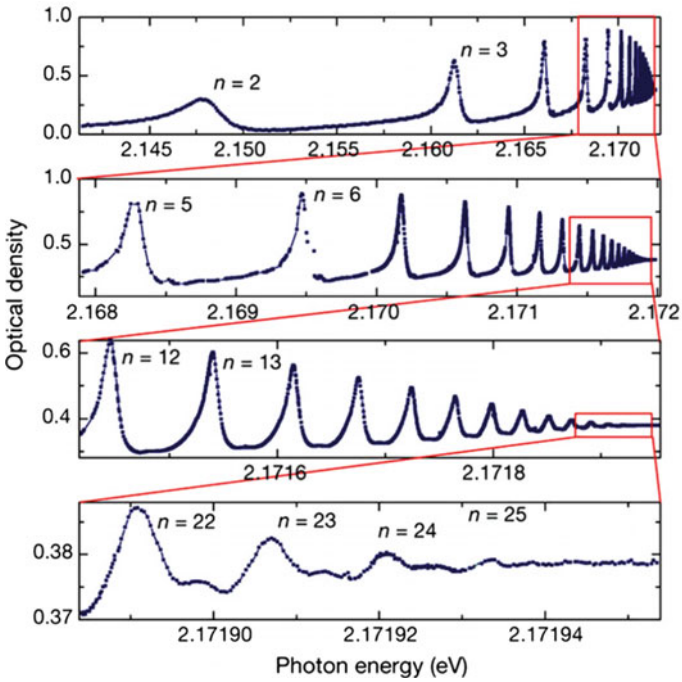


Fig. 21.26 High-resolution absorption spectra of nP excitons in Cu_2O displaying resonances of Rydberg excitons up to $n_B = 25$ (adapted with permission from [14K1] ©2014 Springer Nature)

mixing of states with high angular momentum [15T1, 16S3, 17S1], deviations in the energy level spectrum [16S2], zero-field splittings and different scaling laws when applying external fields [17H2], mixing of states with different parity and optical access to dark states in applied electric fields [17H3]. The interaction of the giant Rydberg excitons with phonons leads to quantum chaotic behavior in the presence of a magnetic field [16A1].

Further recent topics in Cu_2O exciton spectroscopy include the determination of the absolute luminescence yield as a function of temperature [99J1, 01J1], the observation of the excitonic Lyman series (see Sect. 21.3), the investigation of coupled exciton polaritons—acoustic phonon states (phonoritons) [00H1], the observation of extremely narrow 1S absorption peaks [03D1, 07B1], the investigation of the effective masses of ortho- and para-excitons, which become different and direction dependent due to exchange interaction [03D1, 05F1]—more recently attributed to valence-band characteristics [16S3], or of extremely weak LO-phonon assisted absorption below the exciton resonance deduced from luminescence and generalized Planck's law [95D1, 99J1].

21.2.2 Indirect-Gap Semiconductors

Many semiconductors, some of great technological importance such as Si or Ge, but also GaP or AgBr, have an indirect fundamental bandgap. As a consequence, the lowest free excitons states are situated at $\boldsymbol{K} \neq 0$, as outlined in Sect. 20.2. Due to momentum-conservation, these states cannot couple directly to the radiation field. In both absorption and emission processes a third, momentum conserving particle has to be involved (usually a phonon) unless momentum conservation is relaxed as in the alloy semiconductor $Al_{1-x}Ga_xAs$ above the transition composition [96K1]. At low temperature only photon absorption with the creation (i.e., emission) of one or more phonons is possible. At higher temperatures, when the relevant phonon states are populated thermally with finite probability, absorption of light quanta is possible with both phonon emission and absorption.

Figure 21.27a shows schematically the process of absorption accompanied by the creation or emission of a phonon, and in Fig. 21.27b, as an example, the absorption spectrum of GaP at low temperature. The exciton states do not appear as peaks as in direct-gap materials (Fig. 21.8) but instead as onset energies, since the participation of phonons makes the whole density of states available. Thus the absorption spectra are given in the simplest approximation by sums of expressions like

$$\alpha_{\mathrm{em}}(\hbar\omega) \propto \alpha_0[\hbar\omega - E_{\mathrm{ex}}(\boldsymbol{k}_0) + \hbar\omega_{\mathrm{ph}}]^{1/2}(1 + N_{\mathrm{ph}})\,, \tag{21.18a}$$

$$\alpha_{\mathrm{abs}}(\hbar\omega) \propto \alpha_0[\hbar\omega - E_{\mathrm{ex}}(\boldsymbol{k}_0) - \hbar\omega_{\mathrm{ph}}]^{1/2}N_{\mathrm{ph}}\,, \tag{21.18b}$$

where α_0 contains the transition matrix element squared, which is assumed not to depend (or depend only weakly) on the momentum of the phonon. $E_{\mathrm{ex}}(\boldsymbol{k}_0)$ is the exciton energy at the indirect minimum, $\hbar\omega_{\mathrm{ph}}$ the phonon energy for $\boldsymbol{k} \approx \boldsymbol{k}_0$, and N_{ph} is the number for the phonons in a given mode.

The photon energy in the square root term, which describes the density of states of the excitons above $E_{\mathrm{ex}}(\boldsymbol{k}_0)$, has to be chosen so that the argument is positive. For other values of $\hbar\omega_{\mathrm{ph}}$, α_{em} and α_{abs} are zero.

The participation of the phonon makes the transition probability for absorption in indirect-gap semiconductors several orders of magnitude smaller than that for direct, dipole-allowed transitions. Typical values of the "indirect" absorption coefficient are in the range $1\ \mathrm{cm}^{-1} \leq \alpha_{\mathrm{ind}} \leq 10^2\ \mathrm{cm}^{-1}$. These values are so low that no significant structures appear in the reflection spectra.

In Fig. 21.28 we show the onset of the phonon-assisted exciton absorption band of Ge (actually $\alpha^2\ (\hbar\omega)$) and the wavelength modulated spectrum $\Delta\alpha(\hbar\omega)$, exhibiting some structures due to the mass reversal of the exciton, i.e. a change of the curvature, and to the anisotropy of its dispersion close to the L-point (see Sect. 15.2.1).

The absorption process under absorption of phonons is even weaker at low temperatures, i.e., at $k_{\mathrm{B}}T < \hbar\omega_{\mathrm{Phonon}}$. Nevertheless, it was possible to deduce the absorption spectra, e.g., in Si down to $\alpha = 10^{-17}\ \mathrm{cm}^{-1}$ at 90 K from the luminescence spectra, from Kirchhoff's law and the generalized Planck equation [95D1, 95W1].

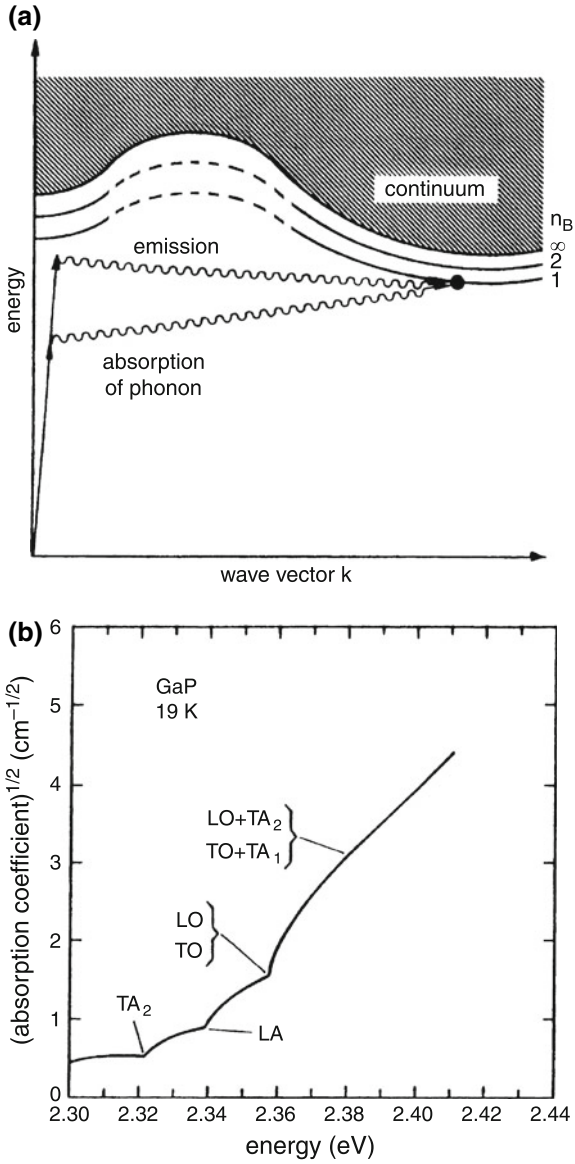


Fig. 21.27 The creation of an exciton in an indirect-gap semiconductor accompanied by creation or annihilation of a momentum-conserving phonon (**a**). The absorption spectrum of GaP (**b**) (reprinted with permission from [70M1] ©1970 Springer Nature)

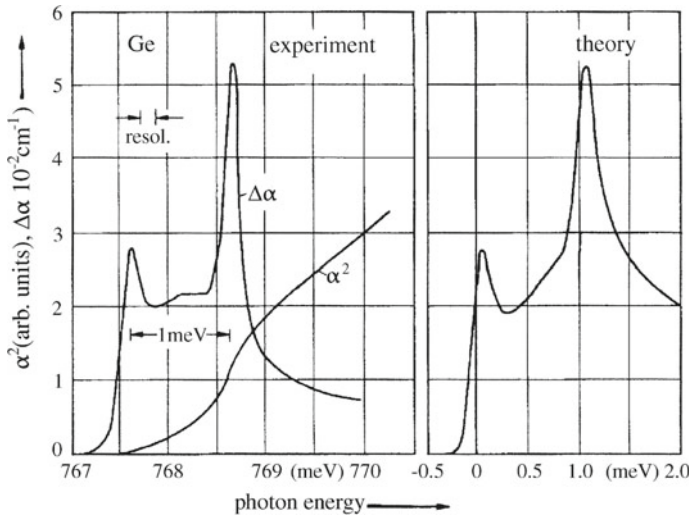


Fig. 21.28 The spectra of the square of the absorption coefficient of the phonon-assisted absorption onset in Ge and its wavelength modulated signal (reprinted with permission from [75F1] ©1975 American Physical Society https://doi.org/10.1103/PhysRevLett.34.1572)

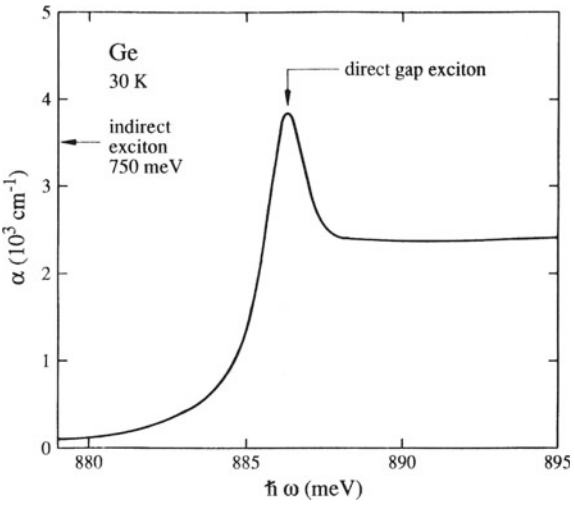


Fig. 21.29 Absorption around the direct exciton in Ge (adapted with permission from [83S2] ©1983 American Physical Society https://doi.org/10.1103/PhysRevLett.51.698)

At higher energies above the indirect exciton, there can also be direct excitons which show absorption structures of the type already discussed in the preceding section. We give an example in Fig. 21.29 for Ge. Since the direct exciton can decay rapidly into lower states, it is strongly damped, preventing the observation of finestructure such as states with $n_B > 1$.

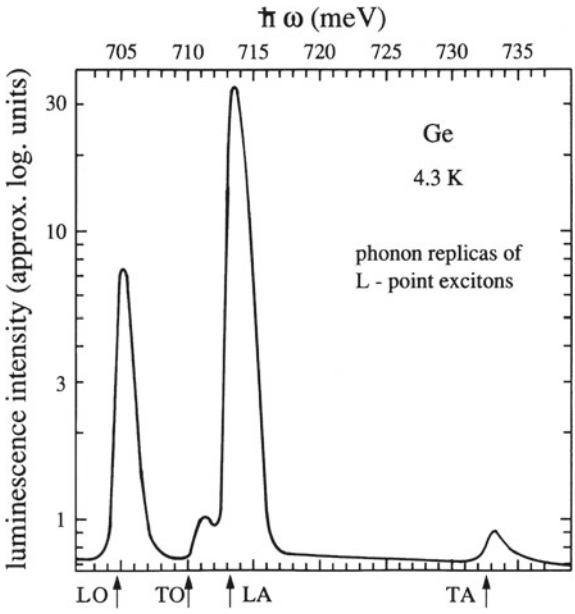


Fig. 21.30 A luminescence spectrum of Ge showing various phonon satellites (according to [76T1])

The luminescence of free excitons in indirect-gap materials involves—as does the absorption—one or more phonons (except for a zero-phonon line in alloys) (Fig. 21.27a). Consequently the exciton-related emission is red-shifted with respect to $E_{ex}(\boldsymbol{k}_0)$. In Fig. 21.30 we show an example for Ge where one sees the contributions of the various phonons.

Due to the participation of a third particle, the luminescence yield of indirect-gap materials is much smaller than that of direct ones. This makes the application of bulk Si in light emitting devices impossible. On the other hand, the lifetime of excitons (or carriers) is rather long in indirect materials and ranges from less than 1 μs to many μs depending on the purity of the material. The lifetime in direct-gap materials is in the (sub-)ns range as will be demonstrated later.

21.3 Intraexcitonic Transitions

Excitons are frequently compared to hydrogen or positronium atoms, essentially with modifications of the masses of the two charge carriers and of the dielectric constant of the embedding medium. We have also used this approach, e.g., in Sects. 20.1 and 20.2. Concerning the spectroscopy of excitons described, e.g., in this chapter so far and of H atoms, there is a crucial difference. If we perform, e.g., absorption spectroscopy on H atoms, the H atoms are already present before we shine light on them. Using the light beam we measure, e.g., the transitions from the 1s state to the n_Bp states ($n_B = 2, 3, \ldots$), i.e., the *Lyman series*.

In a semiconductor, in contrast, there is no exciton before the light is switched on, i.e., there is a *vacuum state* and the incident light creates the excitons, e.g., in the n_Bs states ($n_B = 1, 2, 3, \ldots$). A further difference resulting from the a priori presence of H atoms is the fact that the 1s→2p transitions are usually Doppler broadened, while the excitons can be resonantly created on a well-defined point on their dispersion relation.

The question is now if it is possible to observe the Lyman series also for excitons? The answer is yes! What one needs is a certain density of 1s excitons in the semiconductor created, e.g., by optical pumping from the vacuum into some exciton states (including their continuum). With a probe beam in the suitable spectral range one tries to detect the changes in the transmission induced by the created excitons. This so-called *pump-and-probe* or *differential transmission spectroscopy* is, after two-photon absorption and hyper-Raman scattering (see 21.1.4), a further example of nonlinear optical spectroscopy. We will see more examples for these techniques in the second volume of this textbook.

Alternatively, one may create excitons in their 2P states by some means (e.g., resonant excitation) and measure the luminescence resulting from the transition to the 1S state. We shall see examples and give references for both possibilities in the following.

Early examples of such transitions between excitonic sublevels, also known as *intra-excitonic transitions*, have been given for the indirect gap materials Si and Ge [76T2, 78T1, 88L1]. We give here an example for Si in Fig. 21.31a. The sample is cw pumped by a laser and the *induced absorption* is plotted. The three main features ranging from 10 to 12 meV are attributed to the transitions from the 1S exciton state to various 2P states split by the anisotropy and degeneracy of the band structure of Si at the Δ-point, including the interaction of the carrier spins and the angular momentum of the envelope function.

Though there is a clear signature of these transitions, there are not too many reports of bulk materials for the following reason: The transition energies fall in the range of a few to over 100 meV depending on the exciton binding energy necessitating Fourier spectroscopy or other means of THz spectroscopy. The transition energy should not coincide with some phononic absorption features and, finally, the lifetime of the excitons should be long, so that it is possible to accumulate a sufficiently high density n of excitons under cw pumping, which should exceed, e.g., the density of bound exciton sites (see Sect. 22.1).

Apart from the indirect-gap semiconductors like Si or Ge mentioned above, Cu_2O fulfills the requirements because it has a direct but dipole forbidden band-to-band transition, as detailed in Sect. 21.2, which results in rather long lifetimes [79M2]. Indeed signatures of 1S → 2P transitions have been predicted in [84N1] and reported e.g., in [03F1, 04K1, 06T1], and references therein. However, contributions from defect states, from biphonons or from phonon overtones cannot be ruled out in these data [04J1, 04K2]. The transition from the 1S para exciton to 2P states has been identified [03J1, 04J1, 04K1, 04K3].

We show an example in Fig. 21.31b. Though the total feature looks temptingly similar to the calculated induced absorption spectrum of a Bose condened gas of para

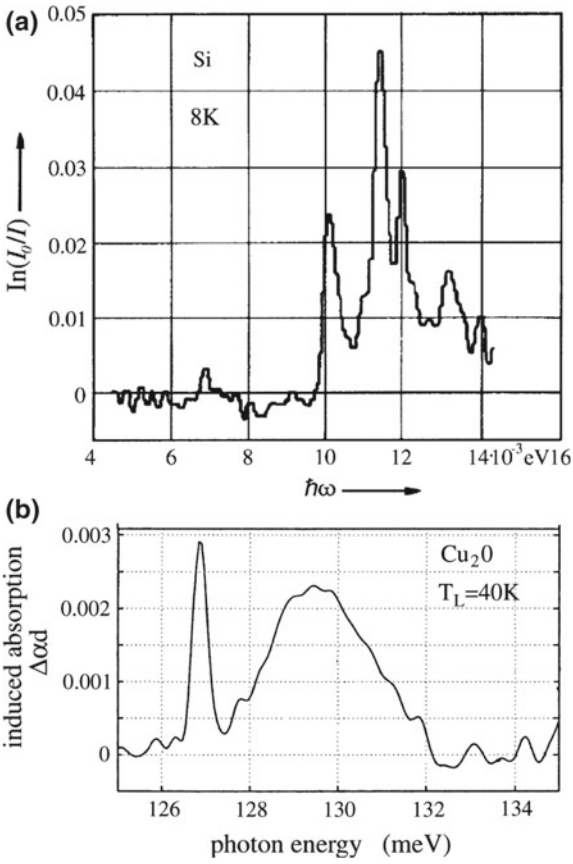


Fig. 21.31 The 1S → 2P transitions in an optically-pumped Si crystal at 8 K (**a**) (reprinted with permission from [78T1] ©1978 Elsevier) and in Cu_2O (**b**) (according to [04J1])

excitons [01J2] (see volume two) a detailed analysis shows, that the peak at 129 meV is due to a 1S para → 2S para transition in a classical Boltzmann exciton gas while the narrow line at 127 meV is probably due to a similar transition in a bound exciton complex (see Sect. 22.1).

Time resolved spectroscopy with THz probe pulses (see volume 2) allows one to observe the 1S → 2P transition in GaAs under resonant exciton creation and to follow the rather slow relaxation of excitons from the continuum states into their 1S ground state in this material and the modifications of the excitation-induced transition if one reaches with increasing pump power an electron–hole plasma [03K1, 05H1, 06K1]. The (stimulated) emission of the 3P → 2S transition by THz spectroscopy has been reported in [06H2, 09L1].

Problems

21.1 Make a plot of the longitudinal-transverse splitting of $n_B = 1$ exciton resonances as a function of the exciton binding energy. Include only semiconductors with direct, dipole-allowed band-to-band transitions. Compare with similar figures in Chap. 20.

21.2 Consider a band-to-band transition in a direct-gap semiconductor neglecting the Coulomb interaction between electron and hole and calculate the absorption spectrum for a dipole-allowed and a dipole-forbidden transition, i.e. for a transition with a matrix element varying linearly with $\boldsymbol{k}$.

21.3 Consider the $n_B = 1$ AΓ_5-polariton resonances in CdS (Fig. 21.3b) and determine for a light beam incident at 45° to the surface ($\boldsymbol{E} \perp \boldsymbol{c}$, $\boldsymbol{k} \perp \boldsymbol{c}$) the length and direction of the wave vectors of the propagating modes in the sample and their phase and group velocities. Select a few characteristic photon energies. Explain the term "spatial dispersion".

21.4 Explain the differences between the concepts of polaritons with spatial dispersion and of birefringence.

21.5 Can the transition energy of a dipole-allowed intra-excitonic transition 1s → 2p coincide with the energy of a dipole-allowed TO phonon?

21.6 Consider Fig. 21.8e for $\boldsymbol{E}$ perpendicular to $\boldsymbol{c}$. In the low temperature limit the A and B excitons are resolved, at 295 K no longer due to their increasing homogeneous broadening.
(a) Deduce the exciton binding energy in ZnO from the 4.2 K spectra and mark then the position of the bandgap at 295 K.
(b) Use equations like a square root absorption edge or an equation for the Tauc regime e.g., from [10K1] to fit the low energy tail of the absorption spectrum. Convince yourself that such a fit works only over a limited range of absorption coefficients $\alpha(\hbar\omega)$ and verify that the extrapolation of this fit to $\alpha(\hbar\omega) = 0$ does give the correct value of the bandgap. Demonstrate also that the maximum of the derivative $\mathrm{d}\alpha(\hbar\omega)/\mathrm{d}\hbar\omega$ does not coincide with the gap.

References

[52H1] M. Hayashi, K. Katsuki, J. Phys. Soc. Jpn. **7**, 599 (1952)
[53U1] F. Urbach, Phys. Rev. **92**, 1324 (1953)
[57E1] R.J. Elliot, Phys. Rev. **108**, 1384 (1957). *Polarons and Excitons*, ed. by C.G. Kupper, G.D. Whitfield (Oliver and Boyd, Edinburg, 1963), p. 269
[57M1] W. Martienssen, J. Phys. Chem. Solids **2**, 257 (1957)
[58D1] D. Dutton, Phys. Rev. **112**, 785 (1958)
[58H1] J.J. Hopfield, Phys. Rev. **112**, 1555 (1958)

[59H1] G. Heiland, E. Mollwo, F. Stöckmann, Solid State Phys. **8**, 191 (1959). See also E. Mollwo, Reichsber. Phys. **1**, 1 (1944)
[60H1] J.J. Hopfield, J. Phys. Chem. Solids **15**, 97 (1960)
[62H1] J.J. Hopfield, D.G. Thomas, Phys. Rev. **128**, 2135 (1962)
[62P1] S.J. Pekar, Sov. Phys. Solid State **4**, 953 (1962)
[62S1] M.D. Sturge, Phys. Rev. **127**, 768 (1962)
[63H1] J.J. Hopfield, D.G. Thomas, Phys. Rev. **132**, 563 (1963)
[64K1] M.N. Kabler, Phys. Rev. A **136**, 1296 (1964)
[64M1] G.D. Mahan, J.J. Hopfield, Phys. Rev. A **135**, 428 (1964)
[65C1] M. Cardona, G. Harbeke, Phys. Rev. A **137**, 1467 (1965)
[65C2] M. Cardona, M. Weinstein, G.A. Wolff, Phys. Rev. A **140**, 633 (1965)
[66G1] D.L. Greenaway, G. Harbeke, J. Phys. Soc. Jpn. **21**(Suppl.), 151 (1966)
[66K1] G. Kuwabara, A. Misu, H. Sasaki, J. Phys. Soc. Jpn. **21**(Suppl.), 148 (1966)
[66R1] J. Ringeissen, Phys. Lett. **22**, 571 (1966)
[67H1] N.J. Harrick, *Internal Reflection Spectroscopy* (Wiley, New Jersey, 1967)
[67K1] M.N. Kabler, D.S. Patterson, Phys. Rev. Lett. **19**, 652 (1967)
[67O1] Y. Onodera, Y. Toyozawa, J. Phys. Soc. Jpn. **22**, 833 (1967)
[67V1] Y.P. Varshni, Physica **34**, 149 (1967)
[68L1] W.Y. Liang, A.D. Yoffe, Phys. Rev. Lett. **20**, 59 (1968)
[68S1] B. Segall, G.D. Mahan, Phys. Rev. **171**, 935 (1968)
[68T1] Y. Toyozawa, J. Hermanson, Phys. Rev. Lett. **21**, 1637 (1968)
[68V1] J. Voigt, F. Spiegelberg, Phys. Status Solidi B **30**, 659 (1968)
[69K1] E.O. Kane, Phys. Rev. **180**, 852 (1969)
[70M1] O. Madelung, *Grundlagen der Halbleiterphysik*, vol. 71, Heidelberger Taschenbücher (Springer, Berlin, 1970)
[70W1] J.P. Walter et al., Phys. Rev. B **1**, 2661 (1970)
[71F1] D. Fröhlich, E. Mohler, P. Wiesner, Phys. Rev. Lett. **26**, 554 (1971)
[71K1] M.V. Kurik, Phys. Status Solidi A **8**, 9 (1971)
[71P1] Y. Petroff, M. Balkanski, Phys. Rev. B **3**, 3299 (1971)
[71S1] H. Sumi, Y. Toyozawa, J. Phys. Soc. Jpn. **31**, 342 (1971)
[71S2] D.D. Sell, P. Lawaetz, Phys. Rev. Lett. **26**, 311–314 (1971)
[72B1] A. Bivas et al., Opt. Commun. **6**, 142 (1972)
[72B2] H.B. Bebb, E.W. Williams, Semicond. Semimet. **8**, 182 (1972)
[72D1] J.D. Dow, D. Redfield, Phys. Rev. **85**, 94 (1972)
[72K1] L. Kalok, J. Treusch, Phys. Status Solidi A **52**, K125 (1972)
[72W1] R.L. Weiher, W.C. Tait, Phys. Rev. B **5**, 623 (1972)
[72W2] E.W. Williams, H.B. Bebb, Semicond. Semimet. **8**, 321 (1972)
[73H1] K. Hümmer, Phys. Status Solidi B **56**, 249 (1973)
[73J1] G.H. Jensen, T. Skettrup, Phys. Status Solidi B **60**, 169 (1973)
[73S1] E. Swenberg, N. Geacintov, Org. Mol. Photophys. **1**, 489 (1973)
[74B1] M. Bettini, S. Suga, R. Hanson, Solid State Commun. **15**, 1885 (1974)
[74S1] S. Suga, T. Koda, Phys. Status Solidi B **66**, 255 (1974)
[74T1] E. Tomzig, R. Helbig, Solid State Commun. **15**, 1513 (1974) and J. Lumin. **14**, 403 (1976)
[75F1] A. Frova et al., Phys. Rev. Lett. **34**, 1572 (1975)
[75H1] P. Hiesinger et al., Phys. Status Solidi B **67**, 641 (1975)
[75K1] C. Klingshirn, Dissertation Universität Erlangen (1975)
[75K2] V.A. Kiselev, B.S. Razbirin, I.N. Uraltsev, Phys. Status Solidi B **72**, 161 (1975)
[75L1] J. Lagois, K. Hümmer, Phys. Status Solidi B **72**, 393 (1975)
[76P1] R.W. Pohl, *Optik und Atomphysik*, 13th edn. (Springer, Berlin, 1976)
[76S1] F. Spiegelberg, E. Gutsche, J. Voigt, Phys. Status Solidi B **77**, 233 (1976)
[76S2] E. Strauss, V. Gerhardt, H. Riederer, J. Lumin. **12/13**, 239 (1976)
[76S3] S. Suga et al., J. Lumin. **12/13**, 109 (1976)
[76T1] G.A. Thomas, M. Capizzi, in *Proceedings of the 13th International Conference on the Physics of Semiconductors*, Rome, ed. by F.G. Fumi (North-Holland, Amsterdam/New York, 1976), p. 915

[76T2] T. Timusk et al., Phys. Rev. B **13**, 3511 (1976)
[77C1] A. Coret, A. Fort, Il Nuovo Cim. **39B**, 544 (1977)
[77V1] H. Venghaus, S. Suga, K. Cho, Phys. Rev. B **16**, 4419 (1977)
[77W1] G. Winterling, E.S. Koteles, M. Cardona, Phys. Rev. Lett. **39**, 1286 (1977)
[78B1] P.D. Bloch, C. Schwab, Phys. Rev. Lett. **41**, 514 (1978)
[78G1] G.G. Grosso et al., Solid State Commun. **25**, 435 (1978)
[78H1] K. Hümmer, Habilitation Thesis, Erlangen (1978)
[78H2] K. Hümmer, P. Gebhardt, Phys. Status Solidi B **85**, 271 (1978)
[78L1] J. Lagois, B. Fischer, Festkörperprobleme/Adv. Solid State Phys. **18**, 197 (1978)
[78S1] W. Stößel, H.J. Wagner, Phys. Stat. Sol. (b) **89**, 403 (1978)
[78T1] T. Timusk et al., Solid State Commun. **25**, 217 (1978)
[78U1] R.G. Ulbrich, C. Weisbuch, Festkörperprobleme/Adv. Solid State Phys. **28**, 217 (1978)
[79E1] W. Ekardt, K. Lösch, D. Bimberg, Phys. Rev. B **20**, 3303 (1979)
[79M1] Y. Masumoto et al., J. Phys. Soc. Jpn. **47**, 1844 (1979)
[79M2] A. Mysyrowicz, D. Hulin, A. Antonetti, Phys. Rev. Lett. **43**, 1123 (1979)
[79P1] J. Puls, J. Voigt, Phys. Status Solidi B **94**, 199 (1979)
[79S1] B. Sermage, G. Fishman, Phys. Rev. Lett. **43**, 1043 (1979)
[79U1] R.G. Ulbrich, G.W. Fehrenbach, Phys. Rev. Lett. **43**, 963 (1979)
[79V1] J. Voigt, F. Spiegelberg, M. Senoner, Phys. Status Solidi B **91**, 189 (1979)
[80B1] I. Broser, M. Rosenzweig, Phys. Rev. B **22**, 2000 (1980)
[80B2] K. Bohnert, G. Schmieder, C. Klingshirn, Phys. Status Solidi B **98**, 175 (1980)
[80B3] I. Broser, M. Rosenzweig, Solid State Commun. **36**, 1027 (1980)
[80D1] W. Dreybrodt et al., Phys. Rev. B **21**, 4692 (1980)
[80F1] M. Fukui, V.C.-Y. So, G.I. Stegeman, Phys. Rev. B **22**, 1010 (1980)
[80M1] I.V. Makarenko, I.N. Uraltsev, V.A. Kiselev, Phys. Status Solidi B **98**, 773 (1980)
[80K1] E.S. Koteles, G. Winterling, Phys. Rev. Lett. **44**, 948 (1980)
[81B1] I. Broser et al., Solid State Commun. **39**, 1209 (1981)
[81C1] K. Cho, Y. Yamane, Solid State Commun. **40**, 121 (1981)
[81I1] T. Itho et al., Solid State Commun. **37**, 925 (1981)
[81J1] W.B. Jackson et al., Appl. Opt. **20**, 1333 (1981)
[81K1] C. Klingshirn, H. Haug, Phys. Rep. **70**, 315 (1981)
[81K2] R. Kuhnert, R. Helbig, K. Hümmer, Phys. Status Solidi B **107**, 83 (1981)
[81K3] R. Kuhnert, R. Helbig, App. Opt. **20**, 4149 (1981)
[81L1] J. Lagois, Phys. Rev. B **23**, 5511 (1981)
[81R1] D.C. Reynolds, T.C. Collins, *Excitons: Their Prooperties and Uses* (Academic, New York, 1981)
[81U1] Ch. Uihlein, D. Fröhlich, R. Kenklies, Phys. Rev. B **23**, 2731 (1981)
[82B1] G. Blattner et al., Phys. Rev. B **25**, 7413 (1982)
[82B2] G. Blattner, private communication
[82H1] I. Hirabayashi, Q. Tokura, T. Kada, J. Phys. Soc. Jpn. **51**, 2934 (1982)
[82I1] M. Itoh, I. Ogura, J. Appl. Phys. **53**, 5140 (1982)
[82J1] N.V. Joshi, B. Vincent, Solid State Commun. **44**, 439 (1982)
[82K1] C. Klingshirn, in *Application of High Magnetic Fields in Semiconductor Physics*, Lecture Notes in Physics, vol. 177 ed. by G. Landwehr (Springer, Berlin, 1982)
[82L1] O. Madelung, U. Rössler (ed.), *Landolt–Börnstein*. New Series, Group III, vol. 17 a to i, 22 a and b, 41 A to D (Springer, Berlin, 1982–2001)
[82M1] V.M. Agranovich, A.A. Maradudin (eds.), *Modern Problems in Condensed Matter Sciences*, (North Holland, Amsterdam, 1982 ff) Surface Polaritons, vol. 1 ed. by V.M. Agranovich, D.L. Mills. Surface Excitations, vol. 9, ed. by V.M. Agranovich, R. Landon
[82O1] Y. Onodera, T. Oshikiri, J. Phys. Soc. Jpn. **49**, 1845 (1980) and ibid. **51**, 2194 (1982)
[82P1] S. Permogorov, in *Excitons*, vol. 2, Modern Problems in Condensed Matter Sciences, ed. by E.I. Rasha, M.D. Sturge (North Holland, Amsterdam, 1982), p. 177
[82R1] M. Rosenzweig, Ph.D. Thesis, TU Berlin (1982)
[82S1] T. Skettrup, Phys. Status Solidi B **109**, 663 (1982)

[82S2] L. Schultheis, J. Lagois, Solid State Commun. **44**, 1557 (1982) and Phys. Rev. B **29**, 6784 (1984)
[83G1] R. Grasser, R. Scharmann, N.A.T.O.A.S.I. Ser, B **88**, 317 (1983)
[83M1] W. Maier, G. Schmieder, C. Klingshirn, Z. Phys. B **50**, 193 (1983)
[83S1] Y. Segawa, Y. Aoyagi, S. Namba, J. Phys. Soc. Jpn. **52**, 3664 (1983)
[83S2] H. Schweizer et al., Phys. Rev. Lett. **51**, 698 (1983)
[84F1] M. Fukui, A. Kamada, O. Tada, J. Phys. Soc. Jpn. **53**, 1185 (1984)
[84L1] M.V. Lebedev et al., JETP Lett. **39**, 366 (1984)
[84M1] M. Matsushita, J. Wicksted, H.Z. Cummins, Phys. Rev. B **29**, 3362 (1984)
[84M2] T. Mita, N. Nagasawa, Solid State Commun. **44**, 1003 (1984)
[84M3] A. Manoogian, J.C. Woolley, Can. J. Phys. **62**, 285 (1984)
[84N1] S. Nikitine, J. Phys. Chem. Solids **45**, 949 (1984)
[84R1] R. Ruppin, Phys. Rev. B **29**, 2232 (1984)
[84S1] T. Shigenari, X.Z. Lu, H.Z. Cummins, Phys. Rev. B **30**, 1962 (1984)
[84W1] J. Wicksted et al., Phys. Rev. B **29**, 3350 (1984)
[85B1] F. Beerwerth, D. Fröhlich, Phys. Rev. Lett. **55**, 2603 (1985)
[85H1] B. Hönerlage et al., Phys. Rep. **124**, 161 (1985)
[85I1] K. Inoue et al., Solid State Commun. **54**, 193 (1985)
[85L1] J.G. Liebler, S. Schmitt-Rink, H. Haug, J. Lumin. **34**, 1 (1985)
[86B1] F. Beerwerth, D. Fröhlich, Phys. Rev. Lett. **57**, 1344 (1986)
[86U1] M. Ueta, H. Kazanki, K. Kobayashi, Y. Toyozawa, E. Hanamura, *Excitonic Processes in Solids*, Springer Series in Solid State Sciences, vol. 60 (Springer, Berlin, 1986)
[87T1] C.W. Tang, S.A. Vanslyke, Appl. Phys. Lett. **51**, 913 (1987)
[88J1] A. Juhl, D. Bimberg, J. Appl. Phys. **64**, 303 (1988)
[88L1] D. Labrie et al., Phys. Rev. Lett. **61**, 1882 (1988)
[90B1] J.H. Burroughes et al., Nature **347**, 539 (1990)
[90K1] H. Kuzmany, *Festkörperspektroskopie* (Springer, Berlin, 1990)
[90L1] J.D. Lambkin et al., Appl. Phys. Lett. **57**, 1986 (1990)
[90R1] S. Rudin, T.L. Reinecke, B. Segall, Phys. Rev. B **42**, 11218 (1990)
[91B1] D. Bimberg, T. Wolf, J. Böhrer, NATO ASI Series B, Advances in nonradiative processes in solids, B. Di Bartolo ed. vol. 249 (Plenum, New York, 1991), p. 529
[91F1] D. Fröhlich et al., Phys. Rev. Lett. **67**, 2343 (1991)
[91O1] T. Ogawa, T. Takagahara, Phys. Rev. B **43**, 14325 (1991)
[91U1] R. Ulbrich, in *Materials Science and Technology*, vol. 4, ed. by W. Schröter (VCH, Weinheim, 1991), p. 65
[93F1] M. Fiebig, D. Fröhlich, Ch. Pahlke-Lerch, Phys. Status Solidi B **177**, 187 (1993)
[93H1] H. Haug, S.W. Koch, *Quantum Theory of the Optical and Electronic Properties of Semiconductors*, 2nd edn. (World Scientific, Singapore, 1993)
[93K1] C. Klingshirn, N.A.T.O.A.S.I. Ser, B **301**, 119 (1993)
[93L1] J.L. Lin, J.P. Wolfe, Phys. Rev. Lett. **71**, 1222 (1993)
[93M1] F.M. Mirabella (ed.), *Internal Reflection Spectroscopy: Theory and Applications* (Marcel Dekker Inc., New York, 1993)
[93M2] Y. Ma et al., Phys. Rev. Lett. **71**, 3725 (1993)
[93P1] N. Peyghambarian, S.W. Koch, A. Mysyrowicz, *Introduction to Semiconductor Optics* (Prentice Hall, Englewood Cliffs, 1993)
[93V1] M. Vening, D.J. Dunstan, K.P. Homewood, Phys. Rev. B **48**, 2412 (1993)
[94F1] D. Fröhlich in *Nonlinear Spectroscopy of Solids: Advances and Application* (1993). NATO ASI Series B, vol. 339 (Plenum Press, New York, 1994), p. 289
[94F2] F. Fuchs et al., Superlatt. Microstruct. **16**, 35 (1994)
[94H1] E.I. Haskal, Y. Zhang, P.E. Burrows, S.R. Forrest, Chem. Phys. Lett. **219** 325 (1994); Phys. Rev. B **51**, 4449 (1995)
[94J1] C. Janowitz et al., Phys. Rev. B **50**, 2181 (1994)
[94W1] M. Watanabe, T. Hayashi, J. Phys. Soc. Jpn. **63**, 785 (1994)
[95B1] F. Bassani, G. Czajkowski, A. Terdicucci, Z. Phys. B **98**, 39 (1995)

[95D1] E. Daub, P. Würfel, Phys. Rev. Lett. **74**, 1020 (1995)
[95D2] G.E. van Dorssen et al., Rev. Sci. Instrum. **66**, 1480 (1995)
[95S1] H. Schneider, K. Köhler, Phys. Rev. B **52**, R14364 (1995)
[95W1] P. Würfel, S. Finkbeiner, E. Daub, Appl. Phys. A **60**, 67 (1995)
[96B1] V. Bulovic et al., Chem. Phys. **210**, 1 (1996)
[96K1] H. Kalt, *Optical Properties of IIIV Semiconductors: The Influence of Multi-Valley Band Structures*, Springer Series in Solid-State Sciences, vol. 120 (Springer, Berlin, 1996)
[96M1] A. Mysyrowicz et al., Phys. Rev. Lett. **77**, 896 (1996)
[96N1] U. Neukirch et al., Phys. Status Solidi B **196**, 473 (1996)
[96T1] S. Tanaka, Y. Kayanuma, Solid State Commun. **100**, 77 (1996)
[97G1] N. Grigorchuk, Phys. Lett. A **231**, 245 (1997)
[97J1] A. Jolk, Diplom Thesis, Universität Karlsruhe (TH) (1997)
[97N1] S. Nüsse et al., Phys. Rev. B **55**, 4620 (1997)
[97N2] U. Neukirch et al., Phys. Rev. B **55**, 15408 (1997) and ibid. **57**, 9208 (1998)
[97S1] M.Y. Shen et al., Phys. Rev. B **56**, 13 066 (1997)
[97W1] J. Wrzensinski, D. Fröhlich, Phys. Rev. B **56**, 13087 (1997)
[97W2] R. Westphäling et al., J. Lumin. **72–74**, 980 (1997)
[98J1] A. Jolk, C. Klingshirn, Phys. Stat. Sol. (b) **206**, 841 (1998)
[98K1] M. Kira, F. Jahnke, S.W. Koch, Phys. Rev. Lett. **81**, 3263 (1998)
[98H1] K. Henneberger, Phys. Rev. Lett. **80**, 2889 (1998)
[98W1] R. Westphäling et al., J. Appl. Phys. **84**, 6871 (1998); and J. Cryst. Growth **184/185**, 1072 (1998)
[99A1] A. Agui et al., Phys. Rev. B **59**, 10792 (1999)
[99D1] J.S. Dodge et al., Phys. Rev. Lett. **83**, 4650 (1999)
[99H1] L. Hanke, D. Fröhlich, A.L. Ivanov, P.B. Littlewood, H. Stolz, Phys. Rev. Lett. **83**, 4365 (1999)
[99J1] A. Jolk et al., in *Proceedings of the 24th International Conference on the Physics of Semiconductors*, Jerusalem, 1998, CD, File II C4 (1999)
[99P1] R. Pässler, Phys. Status Solidi B **216**, 975 (1999)
[99R1] A.V. Rodina et al., Phys. Status Solidi B **216**, 21 (1999)
[00H1] L. Hanke et al., Phys. Status Solidi B **221**, 287 (2000)
[00T1] J. Tignon et al., Phys. Rev. Lett. **84**, 3382 (2000)
[01J1] M. Jörger et al., Phys. Rev. B **64**, 113204 (2001)
[01J2] K. Johnsen, G.M. Kavoulakis, Phys. Rev. Lett. **86**, 858 (2001)
[01N1] J.S. Nägerl et al., Phys. Rev. B **63**, 235202 (2001)
[01R1] A.V. Rodina et al., Phys. Rev. B **64**, 115204 (2001)
[01S1] H.C. Schneider et al., Phys. Rev. B **63**, 045202 (2001)
[02L1] R. Lécuiller et al., Phys. Rev. B **66**, 125205 (2002)
[03C1] S.F. Chichibu et al., J. Appl. Phys. **93**, 756 (2003)
[03D1] G. Dasbach et al., Phys. Status Solidi B **238**, 541 (2003) and Phys. Rev. Lett. **91**, 107401 (2003)
[03F1] T. Fleck, M. Schmidt, C. Klingshirn, Phys. Status Solidi (a) **198**, 248 (2003)
[03J1] M. Jörger et al., Phys. Status Solidi B **238**, 470 (2003)
[03K1] R.A. Kaindl et al., Phys. Status Solidi B **238**, 451 (2003)
[03K2] M. Knupfer, Appl. Phys. A **77**, 623 (2003)
[03Z1] H. Zhao, H. Kalt, Phys. Rev. B **68**, 125309 (2003)
[04G1] S. Glutsch, *Excitons in Low-Dimensional Semiconductors* (Springer, Berlin, 2004)
[04J1] M. Jörger, Ph.D. Thesis, Karlsruhe (2004); J. Lumin. **112**, 21 (2005)
[04K1] M. Kuwata-Gonokami et al., J. Phys. Soc. Jpn. **73**, 1065 (2004) and Solid State Commun. **134**, 127 (2005)
[04K2] K. Karpinska, P.H.M. van Loosdrecht, I.P. Handayani, Solid state Commun. **134**, (2005); *EXCON 04*, Cracow, 2004; J. Lumin. **112**, 17 (2005)
[04K3] C. Klingshirn et al., Solid State Commun. **134**, 155 (2005)
[04K4] T. Koida et al., Appl. Phys. Lett. **84**, 1079 (2004)

[04S1] S. Schumacher et al., Phys. Rev. B **70**, 235340 (2004)
[04Y1] X. Yang, D. Neher, D. Hertel, T. Daubler, Adv. Mater. **16**, 161 (2004)
[05F1] D. Fröhlich et al., Solid State Commun. **134**, 139 (2005)
[05G1] J. Grabowska et al., Phys. Rev. B **71**, 115439 (2005)
[05H1] R. Huber et al., Phys. Rev. B **72**, 161314 (2005)
[06H1] R. Hauschild et al., Phys. Status Solidi C **3**, 976 (2006)
[06H2] R. Huber et al., Phys. Rev. Lett. **96**, 017402 (2006)
[06K1] S.W. Koch et al., Nat. Mater. **5**, 523 (2006)
[06K2] J. Kasprzak et al., Nature **443**, 409 (2006)
[06T1] T. Tayagaki et al., Phys. Rev. B **74**, 245127 (2006)
[07B1] J. Brandt et al., Phys. Rev. Lett. **99**, 217403 (2007)
[07K1] C. Klingshirn, Phys. Status Solidi B **244**, 3027 (2007)
[07K2] C. Klingshirn et al., Phys. Rev. B **75**, 115203 (2007)
[08H1] M. Hauser et al., Appl. Phys. Lett. **92**, 211105 (2008)
[08L1] K.G. Lagoudakis et al., Nat. Phys. **4**, 706 (2008)
[09K1] A. Köhler, H. Pässler, Mater. Sci. Eng. R **66**, 71 (2009)
[09L1] T. Leinß et al., Phys. Status Solidi C **6**, 156 (2009)
[10F1] M. Fox, *Optical Properties of Solids*, 2nd edn. (Oxford University Press, Oxford, 2010)
[10K1] C. Klingshirn, Phys. Status Solidi B **247**, 1424 (2010)
[10K2] C. Klingshirn, B.K. Meyer, A. Waag, A. Hoffmann, J. Geurts, *Zinc Oxide: From Fundamental Properties Towards Novel Applications*, Springer Series in Materials Science, vol. 120 (Springer, Heidelberg, 2010)
[10W1] E. Wertz et al., Nat. Phys. **6**, 860 (2010)
[10T1] S. Tanaka et al., Phys. Status Solidi A **207**, 1474 (2010)
[10Y1] P.Y. Yu, M. Cardona, *Fundamentals of Semiconductors*, 4th edn. (Springer, Berlin, 2010)
[11G1] H.M. Gibbs, G. Khitrova, S.W. Koch, Nat. Photonics **5**, 275 (2011)
[12B1] W. Brütting, C. Adachi (eds.), *Physics of Organic Semiconductors*, (Wiley-VCH, Weinheim, 2012)
[12M1] M. Milosevic, *Internal Reflection and ATR Spectroscopy* (Wiley, New Jersey, 2012)
[14C1] I.E. Castelli et al., APL Mater. **2**, 081514 (2014)
[14G1] M. Grätzel, Nat. Mater. **13**, 838 (2014)
[14K1] T. Kazimierczuk, D. Fröhlich, S. Scheel, H. Stolz, M. Bayer, Nature **514**, 343 (2014)
[14S1] M. Saba et al., Nat. Commun. **5**, 5049 (2014)
[14S2] C, Sekine et al., Sci. Technol. Adv. Mater. **15**, 034203 (2014)
[15M1] A. Miyata et al., Nat. Phys. **11**, 582 (2015)
[15T1] J. Thewes et al., Phys. Rev. Lett. **115**, 027402 (2015)
[16A1] M. Aßmann, J. Thewes, D. Fröhlich, M. Bayer, Nat. Mater. **15**, 741 (2016)
[16G1] K. Galkowski et al., Energy Environ. Sci. **9**, 962 (2016)
[16S1] B.R. Sutherland, E.H. Sargent, Nat. Photonics **10**, 295 (2016)
[16S2] F. Schöne et al., Phys. Rev. B **93**, 075203 (2016)
[16S3] F. Schweiner, J. Main, M. Feldmaier, G. Wunner, C. Uihlein, Phys. Rev. B **93**, 195203 (2016)
[16V1] S.A. Veldhuis et al., Adv. Mater. **28**, 6804 (2016)
[17H1] E.M. Hutter et al., Nat. Mater. **16**, 115 (2017)
[17H2] J. Heckötter et al., Phys. Rev. B **96**, 125142 (2017)
[17H3] J. Heckötter et al., Phys. Rev. B **95**, 035210 (2017)
[17M1] J.E. Moser, Nat. Mater. **16**, 4 (2017)
[17S1] F. Schweiner et al., Phys. Rev B **95**, 035202 (2017)
[17Y1] Z. Yang et al., J. Phys. Chem. Lett. **8**, 1851 (2017)
[18R1] T. Ruf et al., Appl. Phys. Lett. **112**, 083902 (2018). https://doi.org/10.1063/1.5017943
[18R1] T. Ruf et al., in *IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC)*, (Waikoloa, HI, 2018)

Chapter 22
Optical Properties of Bound and Localized Excitons

Abstract In the previous chapter we discussed mainly the optical properties of intrinsic free excitons. Here we consider excitons bound to various types of defects and how they are observed in optical spectra (mainly in luminescence). Excitons can also get trapped in localized states in potential fluctuations of disordered bulk materials. We will discuss how this localization affects the optical spectra and present different approaches for the modeling of localization properties. Many of these aspects are also relevant for the structures of reduced dimensionality presented in the next chapter.

22.1 Bound-Exciton and Multi-exciton Complexes

We will start with the description of various bound exciton complexes and elaborate their signatures in optical spectra.

22.1.1 Binding of Excitons to Defects

We have already introduced point defects in Sect. 18.1. Some of these defects can bind excitons resulting in *bound-exciton complexes (BEC)*. In Fig. 22.1 we illustrate excitons bound to an ionized donor (D^+X), a neutral donor (D^0X), and a neutral acceptor (A^0X). The absorption and emission lines of A^0X, D^0X and D^+X are often labelled I_1, I_2 and I_3 lines, respectively.

Excitons are usually not bound to ionized acceptors, since the formation of a neutral acceptor and a free electron is frequently energetically more favorable. This is due to the fact that the hole mass is usually considerably heavier than the reduced mass of electron and hole [10K1, 10K2, 10M1].

The binding energy E^{b} is defined as the energy distance from the lowest free-exciton state at $\mathbf{k} = 0$ to the energy of the complex. It usually increases according to

$$E^{\mathrm{b}}_{\mathrm{D^+X}} < E^{\mathrm{b}}_{\mathrm{D^0X}} < E^{\mathrm{b}}_{\mathrm{A^0X}} \,. \tag{22.1}$$

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_22

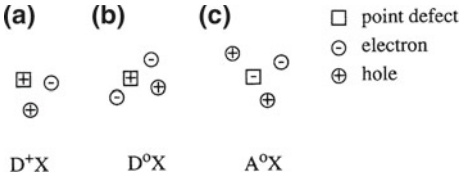


Fig. 22.1 Visualization of an exciton bound to an ionized donor (**a**), a neutral donor (**b**), and a neutral acceptor (**c**)

See for a more detailed discussion of this relation [10M1] and references therein.

There is a rule of thumb, known as *Hayne's rule*, which relates the binding energy of the exciton to the neutral complex with the binding of the additional carrier to the point defect. For the D^0X complex, for example, it states that the ratio of $E^{\mathrm{b}}_{\mathrm{D^0X}}$ to the binding energy of the electron to the donor $E^{\mathrm{b}}_{\mathrm{D}}$ is a constant, depending only on material parameters such as the effective masses:

$$E^{\mathrm{b}}_{\mathrm{D^0X}}/E^{\mathrm{b}}_{\mathrm{D}} = \mathrm{const}\,. \tag{22.2a}$$

The constant is often found to have a value around 0.1–0.2 (for Hayne's rule see [60H1, 77H1, 04M1]). More recently this equation has been expanded by a constant term to [04M1, 10K2]

$$E^{\mathrm{b}}_{\mathrm{D^0X}} = aE^{\mathrm{b}}_{\mathrm{D}} + B \tag{22.2b}$$

The binding energy of the exciton to the complex depends also on the chemical nature of the complex (known as *chemical shift* or *central-cell corrections*) and on the surroundings, leading in high resolution spectroscopy to a splitting of the I_i lines.

Furthermore, bound exciton complexes may have a certain manifold of excited states due to the various mutual arrangements and envelope functions of the two electrons (holes) and the hole (electron) in the D^0X (A^0X) complex [79D1, 81B1]. We give some examples below.

The wavefunctions of excitons bound to shallow centers can be described by a superposition of free exciton wavefunctions in a similar way to that shown in (14.12) for free carriers.

Also isoelectronic traps (such as a Te on a Se site in ZnSe) sometimes form deep centers, with binding energies for excitons exceeding those of neutral acceptors. A long-standing and controversially discussed problem In this context is that of the EL2 luminescence in GaAs. For details including related defects see, e.g., [93S1, 97S1, 98S1].

Donor–acceptor pairs (Fig. 18.1) actually can be considered as poly-centric bound excitons. On the other hand, a single center can, under certain conditions, bind several excitons. The formation of such multi-exciton complexes is especially favored in indirect semiconductors due to the high degeneracy of the multi-valley conduction band and the fourfold degenerate Γ_8^+ valence band [78T1, 89G1].

It is obvious that bound exciton complexes also exist in low-dimensional semiconductor structures. In these systems, the energy of the bound exciton depends in addition on the spatial position of the impurity relative to the barriers. The binding

energy usually decreases if the impurity is located not in the center of the potential well but closer to one of the barriers since the wavefunction is pushed away from the impurity [88F1]. This phenomenon results in an additional inhomogeneous broadening of the absorption and emission lines, which then often merge with the tail-states caused by disorder (see below) and/or with the free-exciton line.

22.1.2 Spectroscopy of Bound Exciton Complexes

Bound exciton complexes do not have any degree of freedom for translational motion. As a consequence BEC often show up in luminescence and absorption spectra as extremely narrow peaks. In ZnO, e.g., linewidths below 0.1 meV have been reported in high quality samples at 1.8 K [65S1, 67S1]. In Fig. 22.2 we give as examples luminescence spectra of ZnO, ZnSe and GaAs. See also Fig. 21.8d.

Figure 22.2a gives an overview of the low-temperature luminescence of a high-quality ZnO single crystal from [76H1, 76T1]. The interpretation of the various groups of emission bands is as follows. FE is the recombination of free exciton polaritons (see also Chap. 21). Then comes a group of bound exciton complexes, the assignment of which has changed in the literature over the years. According to the above references the complexes were attributed with decreasing energy, i.e., with increasing binding energy to the defect, to D^+X, D^0X and A^0X, respectively, according to Fig. 22.1. The neutral acceptors were assumed to be, e.g., Na or Li on Zn sites. Later on, a large fraction of the lines D^0X and A^0X has been attributed to excitons bound to neutral acceptors [97G1]. Presently, all of the lines labeled D^0X in the lower line in Fig. 22.2a are assigned to excitons bound to neutral donors based on the analysis of the two electron transitions, while hardly any shallow acceptors have been identified [04M1]. More recently A^0X complexes have been considered again in the spectral range labelled so in Fig. 22.2a [07K1, 10K2, 10K3, 10M1]. For the nitrogen acceptor in ZnO:N see [02Z1, 04R1, 04W1, 09L1, 09R1, 14R1] and references therein.

The still ongoing discussion whether N or Li are shallow or deep acceptors shows, that the investigation of doping is still a relevant problem in ZnO (see, e.g., [10R1, 14R1]), which is directly linked to bound exciton complexes.

A broad PL band occurring at 3.31 eV has been related to excitons bound to structural defects close to the sample surface [07F1, 08S1]. Narrow emission lines (Y lines) in the spectral range between 3.33 and 3.35 eV have been assigned by a combination of spectroscopic methods to deeply bound excitons at extended structural defect complexes which introduce additional donor states in ZnO [11W1].

Independently of the detailed interpretation of the various lines, we see within one group a splitting of the lines caused by the chemical nature of the binding atom—the above mentioned chemical shift or central-cell correction. Similar shifts or splittings can be caused by the presence of other defects in the vicinity of the center binding the exciton.

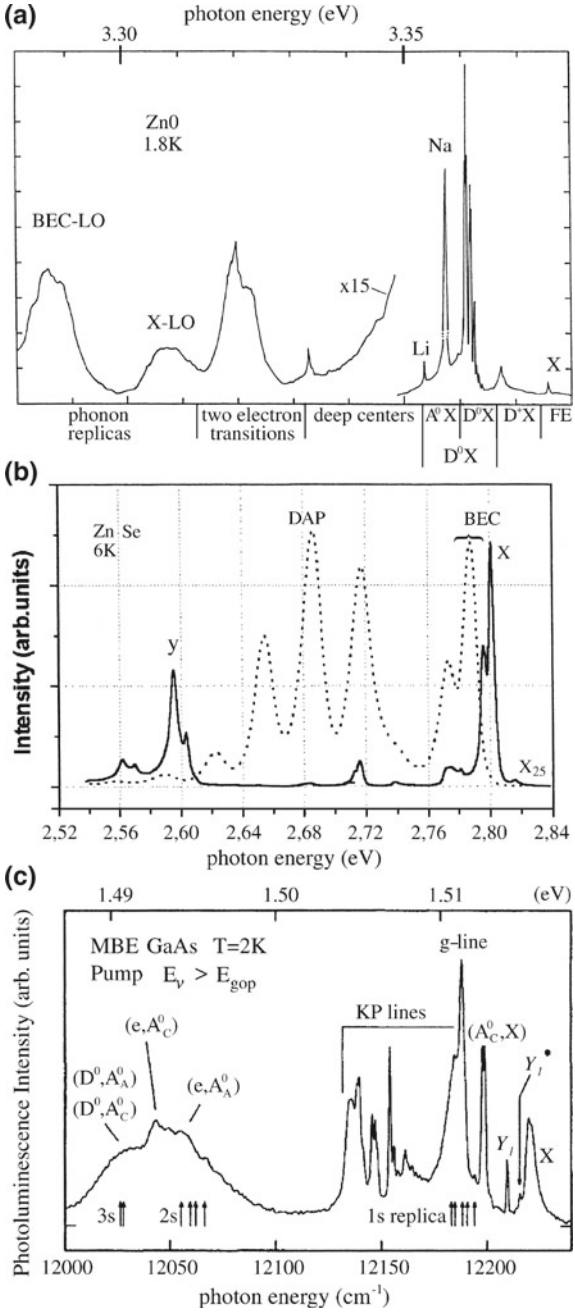


Fig. 22.2 Low-temperature and low-excitation luminescence spectra of ZnO (**a**) [76H1], of two differently grown ZnSe samples (*solid line*: MBE growth with element sources, *dashed line*: with a compound source) (**b**) (adapted with permission from [03D1] ©2003 AIP Publishing https://doi.org/10.1063/1.1568532) and of GaAs (**c**) (reprinted with permission from [91V1] ©1991 American Physical Society https://doi.org/10.1103/PhysRevB.44.13426)

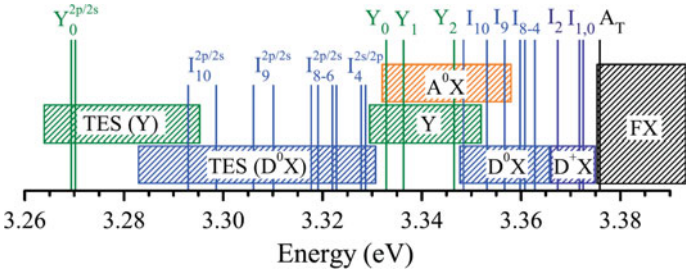


Fig. 22.3 Schematic drawing of the energy ranges of bound-exciton luminescence in ZnO at low temperature. Selected transitions are indicated by vertical lines. The different areas mark the energy range of free excitons (FX), ionized donor bound excitons (D^+X), neutral donor bound excitons (D^0X), acceptor bound excitons (A^0X), deeply bound excitons (Y), and two electron satellites (TES) of shallow and deeply bound excitons in their 2s and 2p states (reprinted with permission from [11W1] ©2011 American Physical Society https://doi.org/10.1103/PhysRevB.84.035313)

Towards lower energies in Fig. 22.2a we see the luminescence of excitons bound to deep centers, possibly deep acceptors. Then follows a range of two electron transitions (see below) and the LO-phonon replica of free and bound excitons, denoted as X-LO and BEC-LO, respectively.

The current assignment of spectral ranges where bound-exciton emission occurs in ZnO to the various species is reviewed in Fig. 22.3 [11W1].

In ZnSe or GaAs we find a similar behavior, as can be seen in Fig. 22.2b, c.

To conclude this part it can be stated that the low-temperature luminescence of high-quality samples is generally dominated by defect- (or localization-)related recombination processes.

Since **k**-conservation is relaxed for BEC due to the lack of translational invariance, BEC can also couple to acoustic phonons, for example via the deformation potential or via the piezoelectric effect. This coupling leads to a so-called *acoustic wing* in the emission or absorption spectra of a BEC. Figure 22.4 shows an example for the I_1 line in CdS. Since only few phonon states are thermally populated at low temperature, the emission manifests itself essentially as a wing on the low energy side according to

$$E_{A^0X} \longrightarrow E_A^0 + \hbar\omega_{1um} + \hbar\Omega_A\,, \tag{22.3a}$$

where $\hbar\Omega_A$ is the energy of the emitted acoustic phonon. The shape of the phonon wing is determined by the energy dependence of the coupling of the BEC to the phonons and by the population of the latter. This applies especially to the ratio of Stokes to anti-Stokes emission. In Fig. 22.4 the influence of the lattice temperature on the emission line shape is illustrated.

Figure 22.5 shows the acoustic-phonon side bands of the I_1 and I_2 BEC in CdS and in addition an absorption spectrum of the same sample. In this special case actually the reabsorption of a rather broad luminescence band appearing under higher excitation was used. The acoustic wing appears in absorption on the high energy side of the zero-phonon line according to

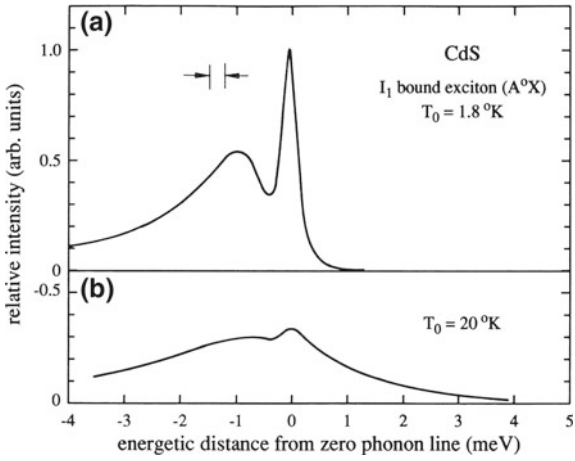


Fig. 22.4 The Stokes and anti-Stokes emission in the vicinity of the I_1 BEC in CdS for two lattice temperatures (adapted with permission from [74S1] ©1974 American Physical Society https://doi.org/10.1103/PhysRevB.10.659)

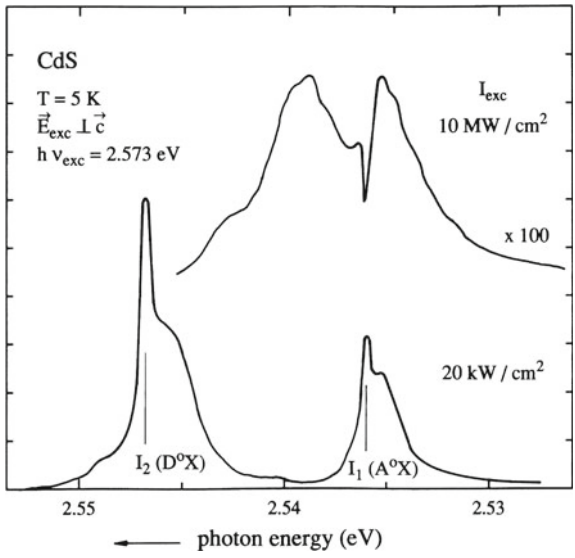


Fig. 22.5 The emission (lower trace) and (re)absorption spectrum of a CdS sample around the I_1 BEC, showing the acoustic wing [79S1]

$$E_{A^0} + \hbar\omega_{abs} \longrightarrow E_{A^0X} + \hbar\Omega_A \,. \tag{22.3b}$$

Another recombination process leaves the point defect in an excited state, this means that the donor electron of a D^0X complex is transferred in the recombination process from its ground state into an excited state. The energy conservation for these *two-electron transitions* reads in the simplest approximation for the D^0X complex, and assuming that the binding energy of the donor forms a hydrogen-like series of states described by the main quantum number n_B,

$$\hbar\omega_{lum} = E_{D^0X} - E_{D^0}(1 - n_B^{-2}) \tag{22.4}$$

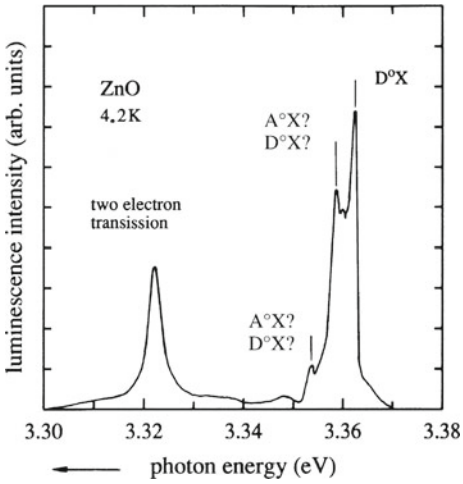


Fig. 22.6 An emission spectrum of a ZnO sample with an especially pronounced two-electron transition (using data published in [75K1]). Concerning the identification of the lower BEC see the discussion in the text

with $E_{D^0} = \mathrm{Ry}\, m_e (m_0\, \varepsilon^2)^{-1}$. In Fig. 22.6 we show an example for ZnO. See also [69R1, 04M1, 04S1, 10M1].

A similar *two-hole transition* may occur for acceptor-bound excitons (A^0X). Corresponding transitions to various 1s, 2s and 3s states are indicated in Fig. 22.2c.

The BEC have a very rich and complex spectrum of excited states. Apart from the quantum numbers known from free excitons n_B, l, m, (see Sect. 20.2), excited states can be created by the relative orientations of the spins of identical carriers (with corresponding parities of the envelope function), or in the A^0X complex, for example, by the participation of holes from deeper valence bands. As an example we give in Fig. 22.7 the photoluminescence excitation spectrum of a bound exciton complex in ZnO (see the discussion above with Fig. 22.2a). The peaks labelled R_2 and R_5 are thought to be due to complexes which involve one or two holes from the B valence band instead of the A valence band. The former is separated from the uppermost A valence band by about 5.4 meV. The resonances R_1, R_3 and R_6 could then be due to the excitation of internal degrees of freedom of the complexes with various hole configurations. Note that the interpretation for R_5 does not work for a D^0X complex, which contains only one hole. More elaborate treatments of this topic are given in [81B1, 86G1, 97G1, 04M1, 10M1].

Many defects like dislocations act as fast nonradiative recombination centers or luminescence killers [00F1]. In some cases they give rise, however, to characteristic luminescence features. The so-called Y-emission line in ZnSe is, e.g., due to the recombination of an exciton bound to the lattice distortion field of a dislocation. An example is given in Fig. 22.2b. The distortions of ZnSe quantum wells which are intersected by a bow-tie like pair of Frank-type stacking faults lead to localization of excitons and an associated pair of bright emission spots in micro-photoluminescence [99L1] or near-field luminescence intensity maps [00F1].

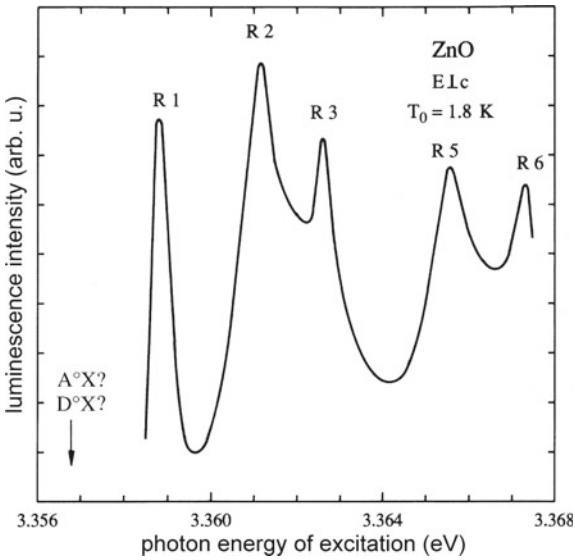


Fig. 22.7 Photoluminescence excitation spectrum of an BEC complex in ZnO (using data from [81B1])

Obviously defects can provide spatially isolated sources of luminescence which should release sequences of photons one at a time. The use of A^0X luminescence in a ZnSe quantum well as a single-photon source has been described in [02S1].

Although we first treat quantum structures in the next section, we want to describe shortly the emission from BEC in quantum wells. Partly their luminescence merges with that of the so-called free, i.e., generally disorder-broadened intrinsic emission lines as is often the case in GaAs/$Al_{1-y}Ga_yAs$ structures. Partly the emission appears as a spectrally separated line. An example is shown in Fig. 22.8 for structures with different quantum-well widths l_z, center-doped with Be acceptors. For further examples see, e.g., [01L1, 10D1]. We recall that for quantum structures, the binding energy of the excitons to the center additionally depends on its position relative to the barriers in addition to the influences of alloy broadening and well-, wire- or dot-width fluctuations etc.

The bandstructure of some indirect-gap semiconductors allows multiple occupation of the conduction band due to the many-valley structure, and of the valence band due to the four-fold degeneracy of the Γ_8 valence-band maximum. In these materials *multiple bound-exciton complexes* can be formed, i.e., BEC which contain one, two, three or even more electron–hole pairs. The emission lines are due to transitions of a complex containing m electron–hole pairs to one with $m - 1$ pairs. In Fig. 22.9 we give as an example a luminescence spectrum of Si:P. The usual ($m = 1$) emission line of the phosphorous D^0X complex dominates. It is followed at lower energies by a series of lines with indices m up to 6. The decay of the higher members of the series is obviously faster than of the lower ones, as expected already from simple reaction kinetics. These multiple BEC can be considered as precursors or nucleation centers for the electron–hole plasma droplets discussed in the second volume of this book.

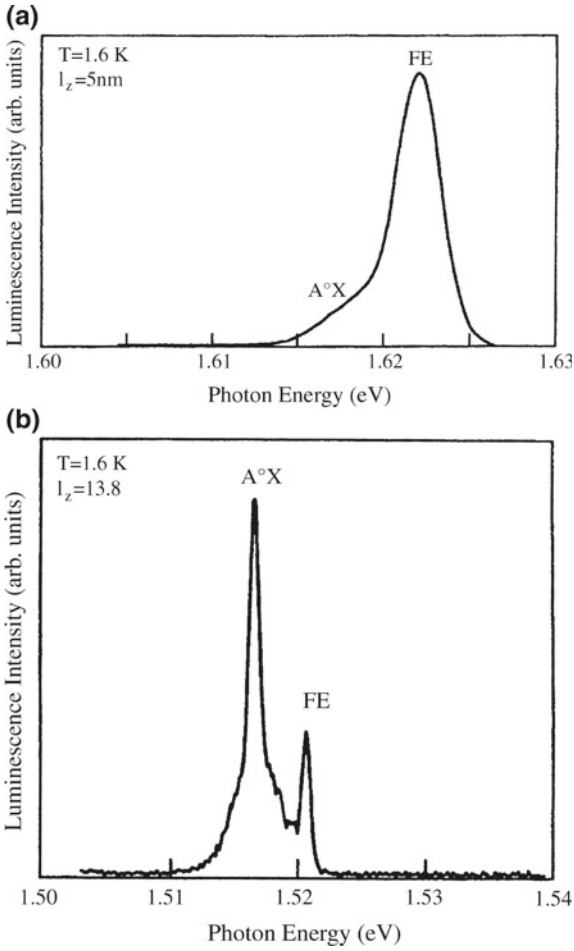


Fig. 22.8 The luminescence of GaAs/$Al_{1-y}Ga_yAs$ structures with two different well widths l_z center-doped with Be acceptors at low temperature showing the emission from localized intrinsic excitons labelled FE and of acceptor bound excitons A^0X (reprinted with permission from [89H1] ©1989 American Physical Society https://doi.org/10.1103/PhysRevB.40.12338)

Another point worth mentioning is that Si is an indirect semiconductor and optical transitions should involve a momentum conserving phonon. However, since the **k**-selection rule is relaxed for localized entities like BEC, the emission lines seen in Fig. 22.9 are actually *zero-phonon lines* which result from recombination without phonon participation.

BECs can be best observed at low lattice temperatures. With increasing T_L the BECs disappear depending on the material, often at temperatures between 20 and 100 K due to thermal dissociation of the exciton from the complex. For illustration compare, e.g., Fig. 21.13 with Fig. 22.2. Below this temperature the halfwidth of the emission lines increase partly linearly and partly quadratically with T_L [65S1, 67S1].

For optimal observation conditions the concentration of the point defects should be low enough ($\lesssim 10^{16}\,cm^{-3}$) to avoid broadening due to interaction between the BEC. Due to this low concentration BEC are usually best observed in lumines-

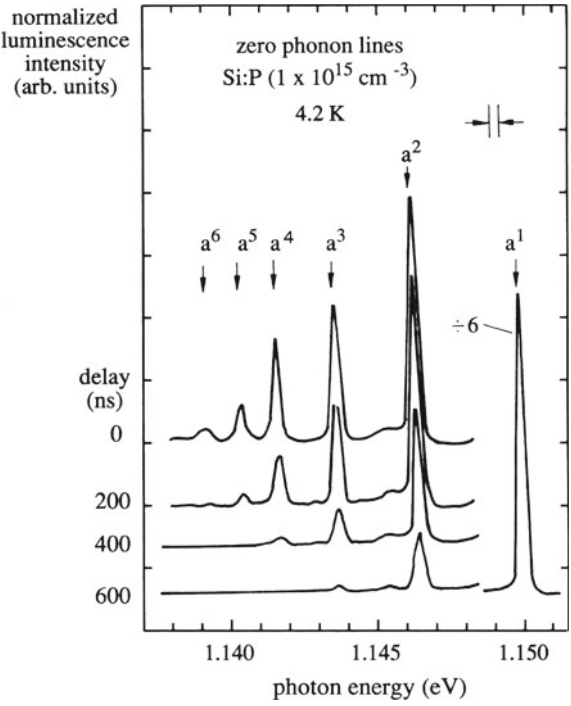


Fig. 22.9 The zero-phonon luminescence spectra of a multiple bound-exciton complex in Si:P after pulsed excitation, normalized to the a_1 line (reprinted with permission from [78T1] ©1978 Elsevier)

cence or photo-luminescence excitation spectroscopy. They show up only weakly in absorption (see, e.g., Fig. 21.8d), and only very rarely do they give rise to reflection structures, simply because the modulation of n and/or κ is too small to result in a significant signal in $R(\omega)$. For an exception to this rule see [78B1].

The occupation of D^0X complexes in GaAs and their manipulation by THz radiation, which brings parts of the donors into excited states, has been monitored by resonant, elastic light scattering in [05A1].

To conclude this section it should be mentioned that BEC are sometimes said to have a *giant oscillator strength* [00R1]. This expression is partly the result of a misconception because the maximum absorption coefficient α_{max} of the free excitons per unit cell was compared with the corresponding quantity of the BEC per defect center. However, in both cases the (Wannier-)excitons cover many unit cells.

For reviews on bound exciton complexes in bulk materials and quantum wells and for some early and more recent observations in various semiconductors see [60H1, 65R1, 65S1, 76D1, 79D1, 82L1, 97G1, 00R1, 01L1, 04M1, 04S1, 05O1, 09I1, 10F1, 10M1, 10S1, 11W1].

22.2 Excitons in Disordered Systems

In our discussion of disordered systems in Sect. 18.2 we saw that disorder leads to the appearance of localized electron and/or hole states (see Fig. 18.8). In a similar way the two-particle complex exciton can be localized in a disordered semiconductor.

If we consider the potential wells and barriers in the valence band e.g., of $CdS_{1-x}Se_x$, in Fig. 18.8 we can envisage two different mechanisms of localization. In very deep potential wells for holes, such a quasi-particle can be localized, and the electron is bound to the localized hole by Coulomb interaction. The other possibility is that there is a wide potential well with dimensions larger than the excitonic Bohr radius. In this case the exciton is localized as a whole. As mentioned already in Sect. 18.2, approaches based on percolation theory are used to describe carrier or *exciton localization*.

In Fig. 22.10 we show schematically the density of localized exciton states for $n_B = 1$. At low energies we start with excitons for which one carrier is localized and the other bound to it by Coulomb attraction. With increasing energy there is a continuous transition to excitons which are localized as a whole, which in turn ends at the transition region to extended exciton states known as the *mobility edge*. Above this region there are the extended exciton states with properties approaching those of free excitons in ordered materials.

In principle, a scheme similar to Fig. 22.10 should also hold for $n_B > 1$. Due to the significant inhomogeneous broadening of the $n_B = 1$ state, however, there is not much chance of identifying higher states of the Rydberg series by optical spectroscopy. Thus usually only the exciton ground states are considered.

As long as the tailing parameter E_0 (18.11) describing the localized excitons is much smaller than E_g, we can use the effective-mass approximation for localized excitons to give the following rules of thumb. Since heavier particles are easier to localize according to (18.12), we will find significant features of localized excitons more frequently in those ionic semiconductors with alloying in the anions that form the valence band like $CdS_{1-x}Se_x$ or $ZnSe_{1-x}Te_x$ than in cation-substituted materials such as $Zn_{1-y}Cd_ySe$.

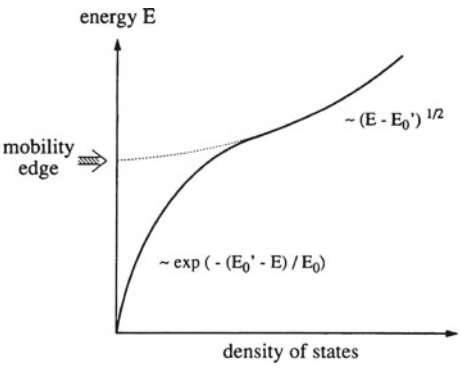


Fig. 22.10 The density of exciton states for $n_B = 1$ in a disordered semiconductor

In semiconductors with more covalent binding, localization occurs both in the conduction and valence bands, as in α-Si or in $Ga_{1-y}Al_yAs$. However, especially in the later example, both the electron and hole masses are relatively low and, as a consequence, it is difficult to localize excitons, i.e., the tailing parameter E_0 and the total number of localized exciton states N_0 are small.

Another aspect of the same feature is the following: Due to the low effective masses mentioned above, the exciton Bohr radius is much larger in $Al_{1-y}Ga_yAs$ than in $CdS_{1-x}Se_x$. This means that the exciton averages over a larger volume, thus diminishing the effective fluctuation and reducing the tail of localized exciton states. As a consequence, in $CdS_{1-x}Se_x$ or $ZnSe_{1-x}Te_x$ for x around 0.1–0.4, the tail of localized exciton states contains roughly 10^{18}–10^{19} states per cm^3, while this number may be one or two orders of magnitude smaller for the $Ga_{1-y}Al_yAs$ system (or for cation-substituted materials). Detailed discussion of the properties of localized excitons is found in [84S1, 85H1, 85H2, 86L1, 89T1, 91C1, 91K1, 92S1, 92S2, 93C1, 93K1, 94R1, 99K1, 99R1, 01R1, 01T1, 01W1, 02R2, 03E1, 04K1, 04K2].

Typical properties of localized states can be deduced from the luminescence spectra. In Fig. 22.11 we show the luminescence of a $CdS_{1-x}Se_x$ crystal under excitation of the extended states and under resonant excitation of the localized ones (see also the spectra in [92P1]). One sees here a rather broad zero-phonon line peaking at 2.21 eV and the first LO-phonon replica around 2.18 eV. The phonon replicas are resolved under resonant excitation into the vibrations of CdS and of CdSe, since in this alloy the phonons are of the persistent-mode type, i.e., they keep approximately the energies corresponding to $x = 0$ and $x = 1$ and contribute to the emission with

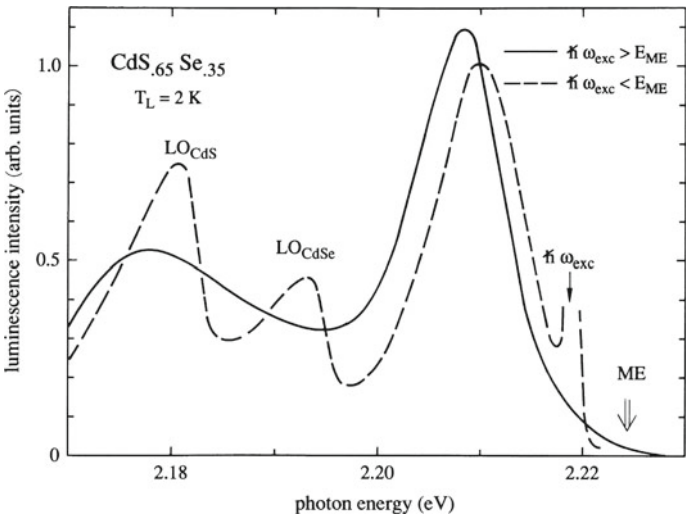


Fig. 22.11 Luminescence spectrum of $CdS_{1-x}Se_x$ under resonant and non-resonant excitation (adapted with permission from [92K2] ©1992 Springer Nature https://doi.org/10.1007/978-1-4615-3044-2_4)

a weight changing with x, while the gap in this case varies continuously with x. See Sects. 13.1 and 18.2, respectively.

The high energy edge of the luminescence band can be identified with the transition region from extended to localized states [92P1, 99R1, 04K1]. This statement can be supported by the following argument: As long as excitons are in extended states, they "scan" the sample and have a good chance of hitting one of the fast non-radiative recombination centers. Once it is trapped in a localized state which does not have such a center within the localization length, the exciton can either recombine radiatively or reach a deeper localized state, e.g., by phonon-assisted tunneling. The probability for the latter process decreases with decreasing energy and density of the localized states, resulting in increasing luminescence yield. Actually at low temperatures the luminescence yield of localized excitons can come close to unity in selected samples [97W1]. These ideas have been confirmed by investigations of the phase-relaxation times and of the polarization memory as will be outlined in the second volume of this book.

A very successful model has been developed for the description of exciton localization based on percolation theory [99K1, 00P1, 04K1]. It has been applied to bulk semiconductor alloys and has been extended for systems of reduced dimensionality. It has also been shown that localization by disorder and binding to shallow defects are competing (or additive, cooperative) effects [10D1]. Percolation of excitons occurs in this model between clusters of atoms of the lower-gap material which form localization sites. If the wave function of excitons in neighboring clusters overlap, they are assumed to form super-clusters. The absorption is governed by ground and excited states of all isolated and super-clusters and their phonon side bands. Low-temperature and low-density luminescence is, in contrast, only emitted from the ground states of isolated and/or superclusters. Taking state filling, hopping and relaxation processes into account as well as electrons and holes in the same or in different (super-)clusters (so-called distant pairs) this model allows one to describe quantitatively the cw absorption and emission spectra as well as the dynamics of the latter after pulsed excitation. The model also describes qualitatively the *S-shaped temperature dependence* of the emission maximum observed in many disordered systems both bulk alloys and quantum wells [01K1, 01R1, 01T1, 01W1, 04K1].

The S-shaped temperature dependence results from the following effects. At higher lattice temperatures defined by $k_B T_L \gtrsim \varepsilon_{loc}$ the majority of the excitons are thermally (re-)excited from the localized into extended states where they behave similarly to free excitons in ordered crystals except for a rather short phase-relaxation time due to disorder scattering. This transition is partly connected to a non-monotonous temperature dependence of the emission maximum and also partly on the FWHM of the band as shown for a $CdS_{1-x}Se_x$ sample in Fig. 22.12, which includes the CdS LO phonon replica. This non-monotonous shift of the emission maximum is frequently called S-shaped behavior in literature, though it more resembles an N-shaped one. Examples of experimental observation and theoretical modelling can be found in [01T1, 02R2, 04K1, 05L1]. Quantum-well systems will be discussed in the next chapter.

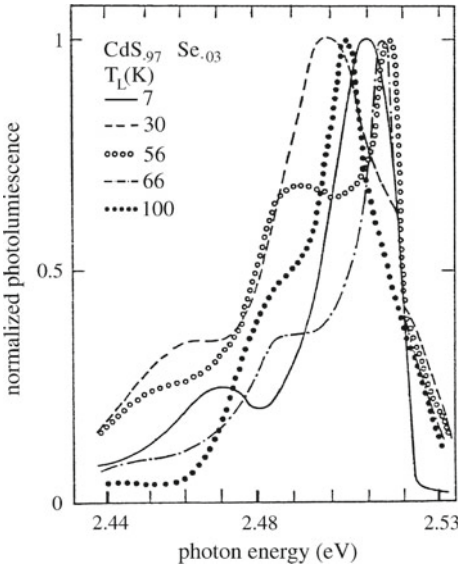


Fig. 22.12 The low excitation photoluminescence spectra of $CdS_{0.97}Se_{0.03}$ for various lattice temperatures T_l (adapted with permission from [88S1] ©1988 John Wiley and Sons)

Similar results as presented for $CdS_{1-x}Se_x$ have also been observed in $ZnSe_{1-x}Te_x$, [92P1] and in $Ga_{1-y}A1_yAs$ [92S2]. In the first case the localization effects are even more pronounced than in $CdS_1 - xSe_x$. Furthermore, self-trapping of excitons seems to occur for $x \approx 0.01$ at single Te atoms or at small Te clusters. In the second case, ε_{loc} is much smaller, partly due to the smaller translational exciton mass which makes localization less probable and partly because of the larger Bohr radius of the exciton which averages over a larger area of the lattice. Also ZnO based alloys like $Zn_{1-y}Mg_yO$ or GaN based ones like $Ga_{1-y}In_yN$ have been investigated with similar results concerning the optical properties of excitons [01P1, 03S1, 10K4, 10M1]. Localization effects occur also in group IV alloys like $Si_{1-x}Ge_x$. For some references see, e.g., [92L1, 92N1, 94T1]. The importance of disorder for Rayleigh scattering and the appearance of speckles is investigated in [04M2, 04Z1].

Actually (quantum-mechanical) level repulsion occurs for states, which are spatially and energetically close lying [03Z1]. This effect should occur for localized states in bulk materials, but is quite difficult to observe experimentally. This is easier to achieve in low-dimensional semiconductor structures.

We will address excitons and their localization in low-dimensional structures in the next chapter. But it is instructive to compare localization in various dimensions already at this point. In quantum wells various types of disorder can contribute to the formation of tails of localized exciton states as already discussed for one-particle states:

- Alloy disorder if the well and/or the barrier material is an alloy. The first case obviously has a stronger influence since the probability of finding the electron and hole in the well of a type-I structure is larger than that of finding it in the barrier.
- Interface roughness, i.e., well-width fluctuations. Usually quantum wells can only be grown with well-width fluctuations of at least one monolayer. These fluctuations of l_z influence the exciton energy via the l_z-dependence of the quantization energy, and via the l_z-dependence of the exciton binding energy (Fig. 9.4). Usually the first effect is the dominant one. For some recent reviews of this topic see [85H1, 93K1, 04K1]. Additional contributions to the localization or binding energy can come from shallow defects like A° or D° in the wells as discussed in Sect. 22.1.

For quantum wells a remarkable evolution has been observed over the last decades concerning the question of exciton localization. In the first euphoric years (in the nineteen seventies and eighties), many authors claimed that the excitonic features in the emission and absorption spectra are all due to free excitons, because the high structural quality of the (MBE-grown) structures. This claims arose despite of the fact that it was well-known that the low-temperature luminescence of even the best parent bulk material was generally dominated by bound exciton complexes and that the low-temperature widths of absorption and emission features were in the quantum wells frequently in the 5–10 meV range, a value far too high to be understood without involving impurity-bound or localized excitons.

Later on, one learned about localization by the various disorder effects in quantum wells mentioned above and began to call the localization sites "quantum dots". Indeed the physical properties of localized carriers or excitons can be comparable to quasi-0D ones. Actually there is no clear and generally accepted criterion to distinguish between a localization site and a quantum island formed by an especially pronounced thickness and/or composition fluctuation of a quantum well. Instead the transition is smooth. It is generally accepted to call, e.g., some $Zn_{1-x}Cd_xSe$ or $Ga_{1-x}In_xAs$ regions with x close to unity and considerably larger than of the embedding quantum well (or wetting layer) between ZnSe or GaAs barriers *quantum islands* or *self-assembled quantum dots*, though the lateral in-plane confinement is usually much smaller than the normal confinement caused by the quantum well (see also [01L1]).

Similar statements and arguments for localization also hold for quantum wires, though localization is less intensively studied in quasi 1d structures. For some examples see, e.g., [91C1, 93C1].

For systems that are confined in all three dimensions, i.e., quantum dots, the question of localization is in principle irrelevant. But carrier and/or excitons can be localized apart from the confining potential of the dot additionally, e.g., at interface states or at localization sites within the dot or in the matrix close to the dot [10S2].

For localization phenomena in various quantum structures see [02R1], for organic alloys see [76K1] and for amorphous semiconductors [03S2].

Problems

22.1. Show by a semiquantitative guess that the A^-X complex is usually unbound.

22.2. For some standard semiconductors, such as Si, Ge, or GaAs, calculate the binding energy of electrons and holes to donors and acceptors, respectively. Find some data for the binding energies of excitons to these complexes and compare the results with Haynes' rule.

22.3. Consider excitons localized in a tailed potential like in Fig. 22.10. Describe qualitatively the photoluminescence lineshape. Why is it useful to plot the intensity of the emission spectrum on a logarithmic scale? What do you expect for very low temperatures?

References

[60H1] J.R. Haynes, Phys. Rev. Lett. **4**, 361 (1960)
[65R1] D.C. Reynolds, C.W. Litton, T.C. Collins, Phys. Rev. **140**, A1726 (1965)
[65S1] Ch. Solbrig, Z. Physik. **184**, 293 (1965)
[67S1] Ch. Solbrig, E. Mollwo, Solid State Commun. **5**, 625 (1967)
[86G1] J. Gutowski, Solid State Commun. **58**, 523 (1986)
[69R1] D.C. Reynolds, C.W. Litton, T.C. Collins, Phys. Rev. **177**, 1161 (1969)
[74S1] J. Shah, R.F. Leheny, W.F. Brinkmann, Phys. Rev. B **10**, 659 (1974)
[75K1] C. Klingshirn, Phys. Status Solidi (b) **71**, 547 (1975)
[76D1] P.J. Dean, D.C. Herbert, D. Bimberg, W.J. Choyke, Phys. Rev. Lett. **37**, 1635 (1976)
[76H1] R. Helbig, Habilitation thesis, University of Erlangen (1976)
[76K1] R. Kapelman, J. Lumin. **12/13**, 775 (1976)
[76T1] E. Tomzig, R. Helbig, J. Lumin. **14**, 403 (1976)
[77H1] B. Hönerlage, U. Schröder, Phys. Rev. B **16**, 3608 (1977)
[78B1] K. Bohnert et al., Solid State Commun. **27**, 295 (1978)
[78T1] M.L.W. Thewalt, Solid State Commun. **25**, 513 (1978)
[79D1] P.J. Dean, D.C. Herbert, in *Excitons*, ed. by K. Cho. Topics in Current Physics, vol. 14 (Springer, Berlin, 1979), p. 55
[79S1] H. Schrey, Doctoral thesis, University of Karlsruhe (TH) (1979)
[81B1] G. Blattner et al., Phys. Status Solidi (b) **107**, 105 (1981)
[82L1] O. Madelung, U. Rössler (ed.), *Landolt–Börnstein*. New Series, Group III, vol. 17a to i, 22a and b, 41A to D (Springer, Berlin, 1982–2001)
[84S1] B.I. Shkolovskii, A.L. Efros, *Electronic Properties of Doped Semiconductors*. Springer Series in Solid-State Science, vol. 45 (Springer, Berlin, 1984)
[85H1] B. Hönerlage et al., Phys. Rep. **124**, 161 (1985)
[85H2] J. Hegarty, M.D. Sturge, J. Opt. Soc. Am. B **2**, 1143 (1985)
[86L1] D.M. Finlayson (ed.), *Localization and Interaction*. Scottish Universities Summer School in Physics, 31st edn. (SUSSP, Edinburgh, 1986)
[88F1] C.Y. Fong, I.P. Batra, S. Ciraci (eds.), *Properties of Impurity States in Superlattice Semiconductors*. NATO ASI Series B, vol. 183 (Plenum, New York, 1988)
[88S1] H.-E. Swoboda et al., Phys. Status Solidi (b) **150**, 749 (1988)
[89G1] J. Gutowski, NATO ASI Ser. B **200**, 139 (1989)
[89H1] P.O. Holtz et al., Phys. Rev. B **40**, 12338 (1989)

[89T1] P.C. Taylor, in *Laser Spectroscopy of Solids II*. Topics in Applied Physics, vol. 65 (Springer, Berlin, 1989), p. 257
[91C1] R. Cingolani et al., Phys. Rev. Lett. **67**, 891 (1991)
[91K1] A.A. Klochikhin, S.G. Ogloblin, Sov. Phys. JETP **73**, 1122 (1991)
[91V1] A. Villemaire, T. Steiner, M.L.W. Thewalt, Phys. Rev. B **44**, 13426 (1991)
[92K2] C. Klingshirn, Excited states in semiconductors, in *Optical Properties of Excited States in Solids*, ed. by B. Di Bartolo, C. Beckwith. NATO ASI Series (Series B: Physics), vol. 301 (Springer, Boston, 1992)
[92L1] L.C. Lenchshyn et al., Appl. Phys. Lett. **60**, 3174 (1992)
[92N1] J.-P. Noël et al., Appl. Phys. Lett. **61**, 690 (1992)
[92P1] S. Permogorov, A. Reznitsky, J. Lumin. **52**, 201 (1992)
[92S1] H. Schwab et al., Phys. Status Solidi (b) **172**, 479 (1992)
[92S2] U. Siegner et al., Phys. Rev. B **46**, 4564 (1992)
[93C1] R. Cingolani, Phys. Scr. **T49B**, 470 (1993)
[93K1] H. Kalt et al., Phys. B **191**, 90 (1993)
[93S1] T.W. Steiner et al., Phys. Rev. B **47**, 1265 (1993)
[94R1] A. Reznitsk et al., Phys. Status Solidi (b) **184**, 159 (1994)
[94T1] K. Terashima et al., Appl. Phys. Lett. **65**, 601 (1994)
[97G1] J. Gutowski, P. Bäume, K. Hauke in *Properties of Wide Gap II–VI Semiconductors*, ed. by R. Bhargava (INSPEC, London, 1997)
[97S1] M.I.N. da Silva et al., J. Appl. Phys. **82**, 3346 (1997)
[97W1] R. Westphäling et al., J. Lumin. **72–74**, 980 (1997)
[98S1] T. Shinagawa, T. Okamure, Jpn. J. Appl. Phys. **37**, 1939 (1998)
[99K1] A. Klochikhin et al., Phys. Rev. B **59**, 12947 (1999)
[99L1] D. Lüerßen et al., Appl. Phys. Lett. **75**, 3944 (1999)
[99R1] A. Reznitsky et al., Phys. Rev. B **59**, 10268 (1999)
[00F1] G. von Freymann et al., Appl. Phys. Lett. **76**, 203 (2000)
[00P1] S. Permogorov et al., J. Cryst. Growth **215**, 1158 (2000)
[00R1] E.I. Rashba, J. Lumin. **87–89**, 1 (2000)
[01K1] E. Kurtz et al., Appl. Phys. Lett. **79**, 1118 (2001)
[01L1] C. Klingshirn (ed.), *Landolt–Börnstein*. New Series, Group III, vol. 34C, Parts 1 and 2 (Springer, Berlin, 2001 and 2004); Part 3, E. Kasper, C. Klingshirn (eds.) (2007)
[01P1] W.I. Park, G.-C. Yi, H.M. Jang, Appl. Phys. Lett. **79**, 2022 (2001)
[01R1] A. Reznitsky et al., in *Proceedings of 9th International Symposium on Nanostructures, NANO 2001*, St. Petersburg, A.F. Ioffe Institute, Russia (2001), p. 538
[01T1] S.A. Tarasenko et al., Semicond. Sci. Technol. **16**, 486 (2001)
[01W1] S. Wachter et al., Phys. Status Solidi (b) **224**, 437 (2001)
[02R1] E. Runge, in *Solid State Physics*, vol. 57, ed. by H. Ehrenreich, F. Spaepen (Academic, San Diege, 2002)
[02R2] A. Reznitsky et al., Phys. Status Solidi C **0**, 280 (2002)
[02S1] S. Strauf et al., Phys. Rev. Lett. **89**, 177403 (2002)
[02Z1] A. Zeuner et al., Phys. Status Solidi B **234**, R7 (2002)
[03D1] M. Dremel et al., J. Appl. Phys. **93**, 6142 (2003)
[03E1] A. Mookerjee (ed.), *Electronic Structure of Alloys, Surfaces and Clusters* (Taylor and Francis, London, 2003)
[03S1] M. Strassburg et al., Phys. Status Solidi C **0**(6), 1835 (2003)
[03S2] J. Singh, K. Simakawa, *Advances in Amorphous Semiconductors* (Taylor and Francis, London, 2003)
[03Z1] R. Zimmermann, E. Runge, V. Savona, Phys. Status Solidi (b) **238**, 478 (2003)
[04K1] A. Klochikhin et al., Phys. Rev. B **69**, 085308 (2004)
[04K2] A. Klochikhin et al., in *27th ICPS*, Flagstaff. AIP Conference Proceedings, vol. 772 (2005), pp. 275, 935
[04M1] B.K. Meyer et al., Phys. Status Solidi (b) **241**, 231 (2004)

[04M2] G. Mannarini, W. Langbein, R. Zimmermann, Phys. Status Solidi C **1**(3), 489 (2004); Phys. Rev. B **69**, 085326 (2004)
[04R1] F. Reuss et al., J. Appl. Phys. **95**, 3385 (2009)
[04S1] R. Sauer, K. Thonke, in *Optics of Semiconductors and Nanostructures*, ed. by H. Kalt, M. Hetterich. Springer Series in Solid State Sciences, vol. 146 (Springer, Berlin, 2004), p. 73
[04W1] D. Wang et al., J. Phys. Condens. Matter **16**, 4635 (2004)
[04Z1] R. Zimmermann, E. Runge, Phys. Rev. B **69**, 155307 (2004)
[05A1] D.G. Allen, M.S. Sherwin, C.R. Stanley, Phys. Rev. B **72**, 035302 (2005)
[05L1] Q. Li et al., Europhys. Lett. **71**, 994 (2005)
[05O1] Ü. Özgür et al., Appl. Phys. **98**, 041301 (2005)
[07F1] J. Fallert et al., J. Appl. Phys. **101**, 073506 (2007)
[07K1] C. Klingshirn, Phys. Status Solidi B **244**, 3027 (2007)
[08S1] M. Schirra et al., Phys. Rev. B **77**, 125215 (2008)
[09I1] T. Ishikawa et al., Phys. B **404**, 4552 (2009)
[09L1] J.L. Lyons, A. Janotti, C.G. van de Walle, Appl. Phys. Lett. **95**, 252105 (2009)
[09R1] I.V. Rogozin, Thin Solid Films **517**, 4318 (2009)
[10D1] C.P. Dietrich et al., New J. Phys. **12**, 033030 (2010)
[10F1] M. Feneberg et al., Phys. Rev. B **82**, 075208 (2010)
[10K1] C. Klingshirn et al., Phys. Status Solidi B **247**, 1424 (2007)
[10K2] C. Klingshirn, B.K. Meyer, A. Waag, A. Hoffmann, J. Geurts, *Zinc Oxide: From Fundamental Properties Towards Novel Applications*. Springer Series in Materials Science, vol. 120 (Springer, Heidelberg, 2010)
[10K3] C. Klingshirn et al., Phys. Status Solidi B **247**, 1424 (2010)
[10K4] C. Klingshirn, in *Zinc Oxide: From Fundamental Properties Towards Novel Applications*. Springer Series in Materials Science, vol. 120 (Springer, Berlin, 2010)
[10M1] B.K. Meyer, in *Zinc Oxide: From Fundamental Properties Towards Novel Applications*. Springer Series in Materials Science, vol. 120 (Springer, Berlin, 2010)
[10R1] C. Rauch et al., J. Appl. Phys. **107**, 24311 (2010)
[10S1] J.W. Sun et al., J. Appl. Phys. **108**, 013503 (2010)
[10S2] A.M. Smith, A.S. Nie, Acc. Chem. Res. **43**, 190 (2010)
[11W1] M.R. Wagner et al., Phys. Rev. B **84**, 35313 (2011)
[14R1] J.G. Reynolds, C.L. Reynolds, Adv. Condens. Matter Phys. **2014**, 457058 (2014)

Chapter 23
Excitons in Low-Dimensional Semiconductor Structures

Abstract Confinement of excitons in low-dimensional structures leads to a strong enhancement of excitonic effect. They have impact on optical properties of these structures up to room temperature even for materials with low excitonic binding in the bulk. We will start in this chapter with the properties of excitons in quasi-2D structures (quantum wells). In particular we will illustrate the dependence of this properties on the confining potential, e.g., its width. The influence of disorder due to well-width fluctuations will be addressed as well as polariton propagation in quantum wells. We will further elaborate the consequences of coupling between neighboring wells—up to the limit of superlattices—on excitonic properties. This includes superlattices with type-II band alignment separating electrons and holes, strained-layer superlattices with tilted bandstructures due to piezoelectric effects, and modulation-doped superlattices. Two-dimensional properties of excitons are best realized in monolayer semiconductors. The excitonic states are here strongly influenced by the unusual bandstructure leading to specific valley-dependent properties. We then proceed to excitons in quantum wires and rods and finally come to quantum dots. We discuss the various confinement regimes of excitons in quantum dots and their consequences on optical spectra. We start with quantum-dot ensembles but we will then describe in detail the fine structure of excitonic states in single InGaAs quantum dots. The later is accessible via micro-photoluminescence or near-field spectroscopy. Another experimental method illustrated in this chapter is photoluminescence excitation spectroscopy (PLE).

23.1 Excitons in (Multiple) Quantum Wells

23.1.1 Excitonic States in (Quasi-)Two-Dimensional Structures

If we consider the exciton again as an effective-mass particle with parabolic dispersion relations, as given by (20.1), we expect a first influence of the dimensionality on the density of states analogous to the situation shown in Fig. 17.1 for every exciton branch $n_B = 1, 2, 3 \ldots$.

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_23

Another effect of the dimensionality manifests itself in the binding energy, the Rydberg series and the oscillator strength. We consider an exciton, for which the motion of electron and hole is restricted to a two-dimensional plane, but the interaction is still a 3D one, i.e., proportional to $e^2/\varepsilon|\boldsymbol{r}_e - \boldsymbol{r}_h|$. This assumption is reasonable since the field lines of the Coulomb field of course penetrate the barrier material (whose dielectric constant contributes to the ε in the strength of the Coulomb force). We find (23.1a) in comparison to the 3D case of (20.1) (see e.g., [93H1]). The dependence of the oscillator strength f on the principal quantum number n_B is given in the limit of (20.2a):

$$3D: E(\boldsymbol{K}, n_B) = E_g - \mathrm{Ry}^* \frac{1}{n_B^2} + \frac{\hbar^2(K_x^2 + K_y^2 + K_z^2)}{2M}, \tag{23.1a}$$

$$f(n_B) \propto \frac{1}{\varepsilon} n_B^{-3}, \tag{23.1b}$$

$$a_B \propto \varepsilon a_B^H n_B, \quad \text{for} \quad n_B = 1, 2, 3 \ldots . \tag{23.1c}$$

For the case of a strict two-dimensional Coulomb interaction one finds with E_Q being the quantization energy:

$$2D: E(\boldsymbol{K}, n_B) = E_g + E_Q - \mathrm{Ry}^* \frac{1}{(n_B - \frac{1}{2})^2} + \frac{\hbar^2(K_x^2 + K_y^2)}{2M}, \tag{23.2a}$$

$$f(n_B) \propto \frac{1}{\varepsilon} \left(n_B - \frac{1}{2}\right)^{-3}, \tag{23.2b}$$

$$a_B \propto \varepsilon\, a_B^H \left(n_B - \frac{1}{2}\right), \quad \text{for} \quad n_B = 1, 2, 3 \ldots . \tag{23.2c}$$

Essentially n_B has to be replaced by $n_B - 1/2$ when going from 3D to 2D systems and the quantization energies E_Q of electrons and holes must be considered. Actually E_Q diverges for confinement in a mathematically strict 2D plane (i.e., $l_z = 0$), see Sect. 17.2. The excitonic Rydberg Ry^* is the same in both cases with the consequence that the binding energy of the 1S exciton is Ry^* in three, and $4\mathrm{Ry}^*$ in two dimensions. The oscillator strength increases and the excitonic Bohr radius decreases when going from three- to two-dimensional systems. Note that the two counteracting terms $+E_Q - \mathrm{Ry}^*/(n_B - 1/2)^2$ in (23.2a) never lead to a red-shift of the exciton state with decreasing dimensionality, since the value in (23.1a) corresponds to the energy minimum already.

The usual realization of quasi-2D excitons is via (M)QW of type I. In this case the motion in the z-direction is quantized, but the width of the quantum well l_z is nonzero. In Fig. 23.1 we show the exciton binding energy for GaAs as a function of l_z for infinitely high barriers. Here the curve reaches $4\mathrm{Ry}^*$ for $l_z = 0$. For finite barrier height the binding energy converges to the value of the barrier material (given by the barrier ε) for $l_z = 0$ passing through a maximum of about 2–3 times Ry^* depending on the material parameters [81M1, 84G1, 88K2].

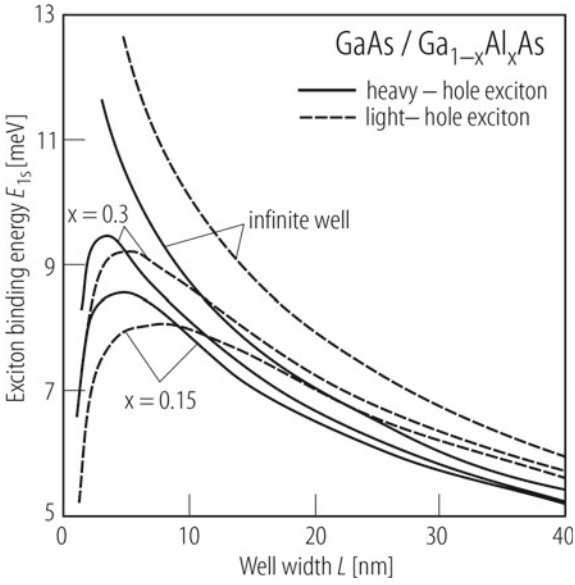


Fig. 23.1 The calculated binding energy of $n_z = 1$hh excitons in GaAs/AlGaAs quantum wells as a function of the well thickness L_z (adapted with permission from [84G1] ©1984 American Physical Society https://doi.org/10.1103/PhysRevB.29.1807)

It is possible to describe the binding energy of the exciton in the quasi two-dimensional case of a QW in terms of an *effective dimensionality* d_{eff} that ranges between three and two and interpolates thus between the limiting cases of (23.1) and (23.2) by [91H1, 92M1]:

$$E_{\text{ex}}^{\text{b}}(1\text{S}) = \frac{\text{Ry}^*}{\left(1 + \frac{d_{\text{eff}}-3}{2}\right)^2}\,, \tag{23.3a}$$

$$\text{with} \quad d_{\text{eff}} = 3 - \exp\left(\frac{-L_{\text{W}}}{2a_\beta}\right) \tag{23.3b}$$

where a_β is the three-dimensional excitonic Bohr radius and L_{W} the width of the quantum well increased by the penetration depth of electron and hole into the barrier.

The increase of the oscillator strength of the 1S exciton comes from the fact that the quantization in the z-direction increases both the overlap between electron and hole and their attraction, which results in turn in a reduction of the two-dimensional Bohr radius.

Also the Sommerfeld factor F, which describes the enhancement of the oscillator strength of the continuum states compared to a situation without election-hole interaction and which arises consequently from the residual electron–hole correlation, depends on the dimensionality [66S1, 91O1]. It reads

$$\text{3D:} \quad F_{\text{3D}} = \frac{\pi}{W^{1/2}} \frac{\text{e}^{\pi/W^{1/2}}}{\sinh(\pi/W^{1/2})} \quad \text{with} \quad W = (E - E_{\text{g}})/\text{Ry}^*\,; \tag{23.4a}$$

$$\text{2D:} \quad F_{\text{2D}} = \frac{\text{e}^{\pi/W^{1/2}}}{\cosh(\pi/W^{1/2})} \quad \text{with} \quad W = (E - E_{\text{g}} + E_{\text{Q}})/\text{Ry}^*\,. \tag{23.4b}$$

In the three-dimensional case it has a square-root singularity at E_g and decreases gradually to unity for $E > E_g$. This one over square root singularity just compensates the square root density of states, resulting above E_g in an (almost) constant absorption spectrum (for examples see Fig. 21.8). In two dimensions it decays only from two to one with increasing energy and in a one-dimensional system it is even below unity above the gap quenching thus the singularity in the (combined) density of states.

The corrections to the simple exciton models which we mentioned in Sect. 20.3 hold partly also in the quasi-2D case. The 2D valence band structure may be even more complex than the 3D one, comparing e.g., Figs. 15.9 and 17.5. However, it should be noted that the most widely investigated (M)QW are based on AlGaAs and InGaAs, which fulfill the inequality (20.7a).

The most striking feature of excitons in quasi-two-dimensional systems and in systems of even lower dimensionality, however, are the facts that exciton series ($n_B = 1, \ldots \infty$) exist for every combination of electron and hole subband. But they partly have small or vanishing oscillator strength and the exciton splits into light- and heavy-hole excitons which results from the corresponding splitting of the valence-band states (Figs. 17.4 and 17.5). To describe the exciton states we thus need more quantum numbers. Apart from the principal quantum number n_B in (23.1) or (23.2), it must be indicated which of the quantized conduction- and valence-band states are involved. The simplest optical interband selection rule is $\Delta n_z = 0$, so that we observe in optical spectra mainly excitons which obey this rule. Finally we must specify whether we are speaking of the light- or the heavy-hole series. Complete information might thus be the $n_z^h = 2\text{hh}, n_z^e = 2, n_B = 1$ exciton state. Usually one uses the abbreviation $n_z = 2\text{hh}$ exciton involving the above selection rule and the fact that excitons in MQW with $n_B > 1$ are usually difficult to resolve due to broadening effects.

Biexcitons also form bound states in quantum wells, wires and dots again in agreement with experiments. The enhancement of the exciton binding energy in structures of reduced dimensionality also shows up in an increasing biexciton binding energy with increasing confinement [83K1, 96S4, 99D2, 05B1, 05B2]. In [96S4] it has been calculated, that the biexciton binding energy in quantum wells should be approximately 22% of the corresponding exciton binding energy independent of the well width. Though doubts have been raised concerning the theoretical model, it was found to be in good agreement with experiment.

23.1.2 Optical Spectroscopy of Excitons in (Multiple) Quantum Wells

Single heterojunctions like in Fig. 17.2 quantize only one type of carriers. The number of publications on optical spectroscopy of excitons in such systems is consequently rather limited (compared to QW and SL). In these cases often a QW is incorporated in the depletion region of the single heterojunction or a high magnetic field is involved.

For a few examples see, e.g., [86A1, 88K3, 89K3, 94R1, 94S1, 95H1] and the references therein. We will focus in the following on single (SQW) and multiple quantum wells (MQW).

The description of optical properties related to excitons in QWs will remain mostly in the weak coupling limit. The main reason is that inhomogeneous broadening due to interface roughness or alloy disorder mask the fine details stemming from the polariton concept see e.g., [84S1, 89K1]. We will come back to the polariton concept at the end of this section.

The most widely investigated groups of MQW are based on material systems GaAs/AlGaAs and InP/InGaAs/InAlAs (see Appendix B) [01K1]. But there has also been considerable interest in II–VI systems [04K1], especially for wide gap materials like structures based on ZnCdMg/SSeTe, ZnMgCd/O and in group III–nitride systems [01K1] like GaAlIn/N and GaInNAs for the blue and IR regions, respectively.

In Fig. 23.2 we show how the absorption spectra of a quantum-well system develop from the bulk material. The layer with $l_z = 400$ nm is almost bulk-like and shows the 1S exciton absorption peak followed by its ionization continuum. Compare this to Fig. 21.8. With decreasing quantum-well width, the excitons of the various subband transitions appear (labeled by $n_z = 1, 2, \ldots$) and the splitting into heavy-hole and light-hole excitons caused by the different, mass-dependent quantization energies.

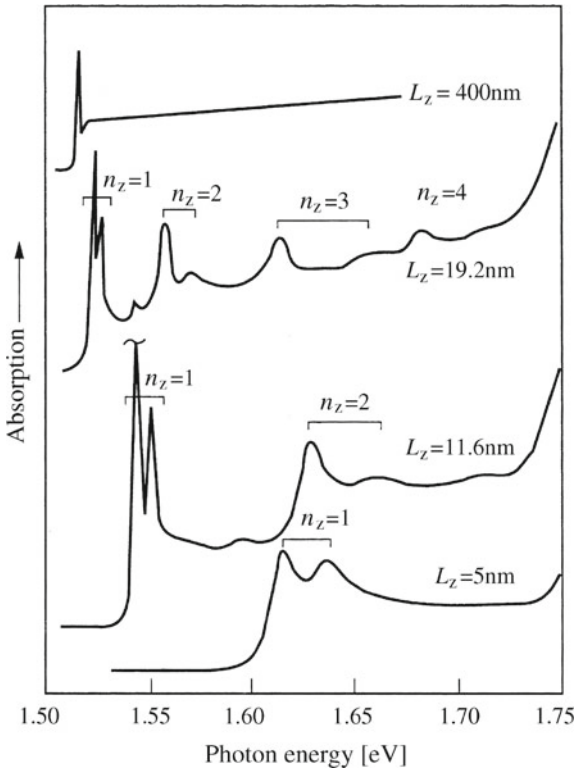


Fig. 23.2 Low temperature absorption spectra of GaAs (quantum) films of various thicknesses (according to data from [83G1], reprinted with permission from [90G1] ©1990 Elsevier)

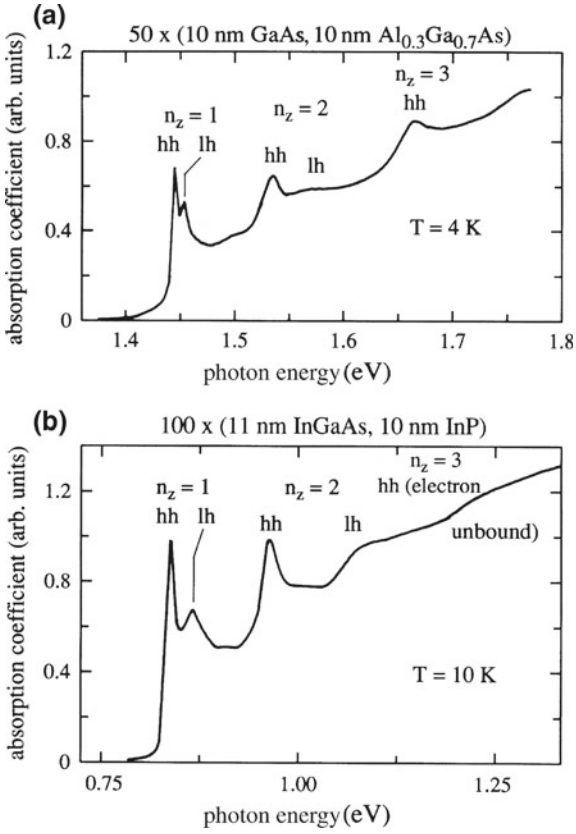


Fig. 23.3 The low temperature absorption spectra of multiple quantum well samples of $GaAs/Al_{1-y}Ga_yAs$ (**a**) (reprinted with permission from [89S1] ©1989 Taylor & Francis Ltd, http://www.tandfonline.com) and of $In_{1-y}Ga_yAs/InP$ (**b**) (according to [01K1])

Some weaker structures appear in between, which are dipole forbidden, but obtain some oscillator strength, e.g., via internal band bending. For further examples see, e.g., [01K1].

In Fig. 23.3a, b we compare the spectra for $GaAs/Al_{1-y}Ga_yAs$ and $In_{1-y}Ga_yAs/InP$ MQWs. We see the exciton and the continuum transitions between the first quantized heavy- and light-hole levels and the first quantized electron state, $n_z = 1$hh and $n_z = 1$lh, respectively. The peaks are due to the $n_B = 1$ state. The higher states $n_B \geqslant 2$ are usually not or only hardly visible, due to their relatively small binding energies and oscillator strengths (see Sect. 23.1.1) and due to homogeneous and inhomogeneous broadening mechanisms. See also below with Fig. 23.4. The absorption coefficient in the continuum states essentially reflects the constant density of states for two-dimensional effective-mass particles. It is not influenced much by the Sommerfeld enhancement because it varies only by a factor two (see (23.4)) unlike three-dimensional systems. The next prominent structures are the $n_z = 2$hh and lh excitons. Higher states are usually less clearly seen, among other reasons because the electron states are often no longer quantized for $n_z > 3$, depending on l_z, m_e and the band offset.

In contrast to the InP system, the GaAs substrate of the AlGaAs MQW is opaque for the exciton resonances. So it has to be removed by selective etching [80P1] for the observation of absorption spectra. In order to avoid this irreversible destruction of the substrate and/or the strain that is often present in the extremely thin remaining film of material on can measure luminescence–excitation spectra. Here the exciton luminescence intensity $I_{\mathrm{lum}}(\hbar\omega_{\mathrm{exc}})$ is detected as a function of the photon energy of excitation for constant incident intensity (or rather constant photon flux density). Similarly, the dependence of the photo current $j_{\mathrm{photo}}(\hbar\omega_{\mathrm{exc}})$ can be measured in a pin diode biased in the blocking direction, which contains the (M)QW in the intrinsic region. The such measured *excitation spectra* are related to, but not identical with, the absorption spectrum $\alpha(\hbar\omega)$ (see Exp. Technique PLE).

In particular, excitation spectroscopy is useful if a single quantum well (SQW) is investigated. The absorption coefficient of two-dimensional excitons in the above-mentioned systems are of the order of $10^4\,\mathrm{cm}^{-1}$. A SQW with $l_z = 10\,\mathrm{nm}$ therefore gives an optical density of only $l_z \cdot \alpha \approx 10\,\mathrm{nm} \cdot 10^4\,\mathrm{cm}^{-1} = 10^{-2}$, which is hardly detectable in a simple transmission experiment. The observation of such a small variation needs either highly developed modulation techniques or the here described types of excitation spectroscopy like photo-luminescence on photo-current excitation spectroscopy.

The advantage of studying SQWs is that fluctuations of l_z between different QWs can be avoided. Furthermore some (opto-)electronic devices contain only a SQW, e.g., some field-effect transistors (MODFET, HEMT).

The additional optical selection rules which arise when the excitons are confined to a QW are rather simple:

$$\Delta n_z = 0\,. \tag{23.5}$$

Transitions with odd Δn_z are forbidden by parity and those for even non-zero Δn_z are forbidden for rectangular shaped QW in the limit of infinitely high barriers. Transitions which violate (23.5) can be observed if there are symmetry-breaking perturbations such as external or internal electric fields. Examples will be given in Sect. 24.2.

The luminescence spectra of excitons in QW are often inhomogeneously broadened and show a Stokes-shift with respect to the absorption. We give two-rather different examples in Fig. 23.4. The effects can be explained by assuming that the luminescence is a superposition of the recombination of excitons in extended states at $\boldsymbol{k}_\parallel \approx 0$ (excitons with larger $\boldsymbol{k}_\parallel$ cannot decay radiatively because of momentum conservation), of excitons localized in tail states due to well-width fluctuations and/or alloy disorder, and finally of excitons bound to impurities such as neutral acceptors or donors. As we saw in Chap. 22, the bound exciton complexes lead to spectrally very narrow emission bands in 3D systems, but in QW their transitions are strongly inhomogeneously broadened because the energy of the complex depends on its positions relative to the barriers. The latter effect can be reduced by deliberately doping only a narrow layer in the center of the well, as shown, e.g., earlier in Fig. 22.8.

In samples of very high quality the mechanisms mentioned above are less pronounced, resulting simultaneously in a reduction of the spectral widths of absorption

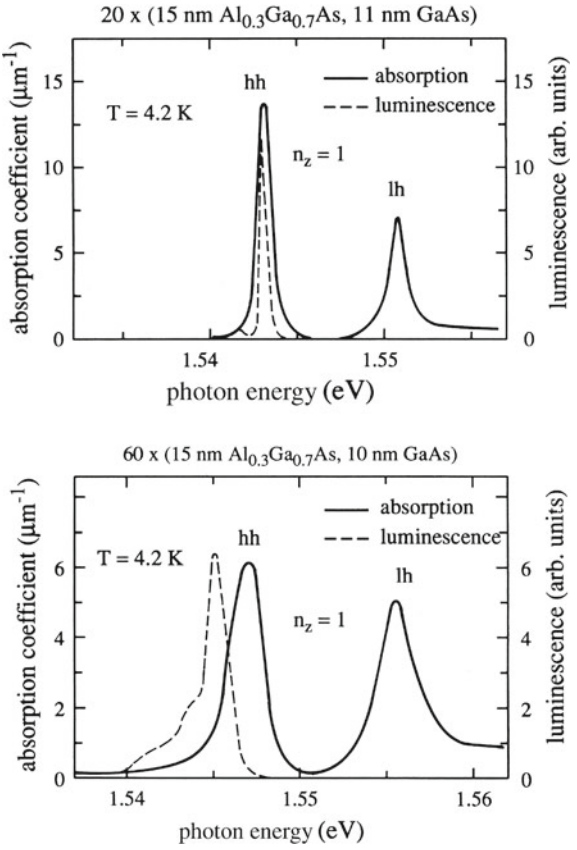


Fig. 23.4 A comparison of absorption and luminescence spectra in two samples of different inhomogeneous broadening (according to [92O2])

and emission bands and of the Stokes shift. Both quantities are related to each other as shown, e.g., in [93Y1, 94G1]. An example is given in Fig. 23.4. The shoulders in the emission on the low energy side can have various origins: fluctuations of l_z by one monolayer, recombination of a bound exciton complex, or recombination of a trion or a biexciton (see Sect. 20.4 for the latter two excitonic species). Due to the weaker coupling of carriers and excitons to optical phonons in the III–V compounds compared to the more ionic II–VI semiconductors, LO-phonon replica are hardly detectable in QWs of the former materials with type-I band alignment.

As an example of a II–VI system we show in Fig. 23.5 the photoluminescence (PL) and photoluminescence excitation (PLE) spectra of a $ZnSe/ZnS_{1-x}Se_x$ MQW sample. One observes in PL the light and heavy hole exciton, a bound-exciton complex (BEC) and their LO phonon replica. The PLE spectrum taken at the position of the hh luminescence shows the lh exciton resonance and above peaks spaced by integer multiples of LO phonons. They indicate that electron–hole pairs first form excitons and then relax by LO-phonon emission in contrast to, e.g., GaAs systems where the carriers relax individually and an exciton is formed only at the end of the

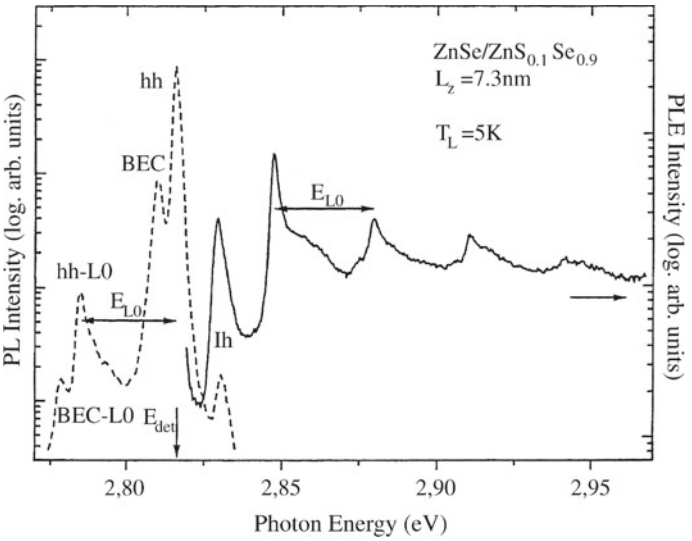


Fig. 23.5 Photoluminescence (PL) (- - - -) and PL-excitation spectra (——) of a $ZnSe/ZnS_{1-x}Se_x$ MQW sample (reprinted with permission from [98K1] ©1998 Elsevier)

LO-phonon emission cascade [93K2, 95B2, 96W3, 96W4]. Concerning the discussion to which extend luminescence at the exciton energy is due to excitons or to a Coulomb correlated electron–hole gas or plasma see [98K1, 02H1, 04C1]. More details on relaxation processes will be given in the second volume of this book.

In this context it is necessary to mention that in the early and euphoric times of quantum-well research (in the early 1980ties), most excitonic low temperature absorption and luminescence bands have been attributed to free excitons, even if their width reached or even exceeded 10 meV. The importance of localization and defect states became clear only later [86A1, 88K3, 90C1, 94F1]. Absorption and emission of free, quasi-twodimensional excitons within the *light cone* $\boldsymbol{k}_\parallel \leq \omega/c$ would result in a linewidth of only one or a few hundred μeV, depending on the effective mass. Even if a homogeneous broadening corresponding to dephasing times on the order of a few ps is added one has to consider a low temperature exciton line width exceeding 1 meV as substantially inhomogeneously broadened.

With the development of *spatially resolved spectroscopy* (micro-photoluminescence μ-PL or scanning nearfield optical microscopy SNOM Sect. 23.5) emission from individual localization sites became visible in the spectra. Some authors even termed these states quasi zero-dimensional quantum dots (see Sect. 17.5). Such localized states have characteristics of quasi-0D states. But they are embedded only in a weakly fluctuating potential landscape with far less confinement than typical for self-assembled quantum dots or colloidal quantum dots [01L1]. Since transfer to neighboring localized sites is rather easily possibly, the term *quantum island* seems to be more appropriate.

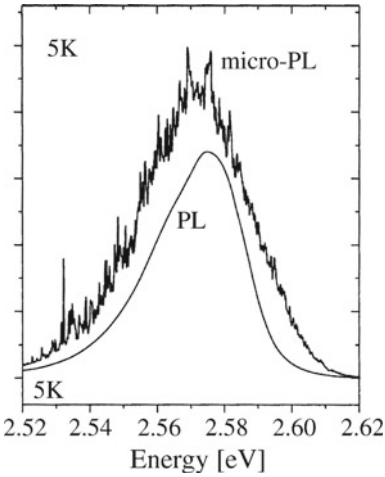


Fig. 23.6 Macro- and micro-photoluminescence spectra of a CdSe/ZnSe quantum well. The nominally deposited amount of CdSe was two monolayers. Due to segregation and diffusion it is spread over about 10 mono-layers and islands of higher CdSe content are formed resulting in the localization sites (reprinted with permission from [01K3] ©2001 Elsevier)

The homogeneous width of excitons localized by well width and alloy fluctuations are at low temperatures often rather narrow (<100 μeV). If only a small ensemble of such localized states is probed, e.g., in luminescence, the localized excitons show up in emission as individual narrow lines while they merge to an inhomogeneously broadened band if their number increases. Examples are found, e.g., in [98P1, 00L1, 01L1]. We give an example in Fig. 23.6, showing a macro- and micro-photoluminescence spectrum (spatial resolution of 1 μm in the latter case) of a CdSe quantum well in ZnSe.

The relaxation and thermalization of the excitons in localized and extended states give rise to a frequently observed non-monotonous dependence of the spectral position of the emission maximum of the macro-PL with temperature. Examples of this non-monotonous variation (called *S-shaped temperature dependence* in the literature) and of the spectral half-width with temperature are given in Fig. 23.7 and in [98P1, 01T1, 01W1, 02R1, 03K2]. The behavior here is very similar to the behavior of localized excitons in bulk alloys, as already shown in Sect. 22.2.

At temperatures below about 40 K the excitons are trapped, after excitation into the continuum states and successive relaxation, randomly to various localization sites. Individual emission lines detected in micro-PL follow the temperature dependence of the band gap. The maximum of the macro-PL, however, shows a stronger red shift for T between 50 and 100 K. In this regime the excitons undergo thermally activated hopping processes to neighboring, deeper localization sites. At even higher temperatures $T > 100$ K scattering into extended states dominates, which have a much higher density of states. Above 150 K the temperature dependence of the PL maximum again roughly follows the band gap. The transition from localized to extended states is frequently accompanied by a non-monotonous behavior of the width of the emission band as shown in Fig. 23.7. For approaches to theoretically model this behavior see [01T1, 04K2].

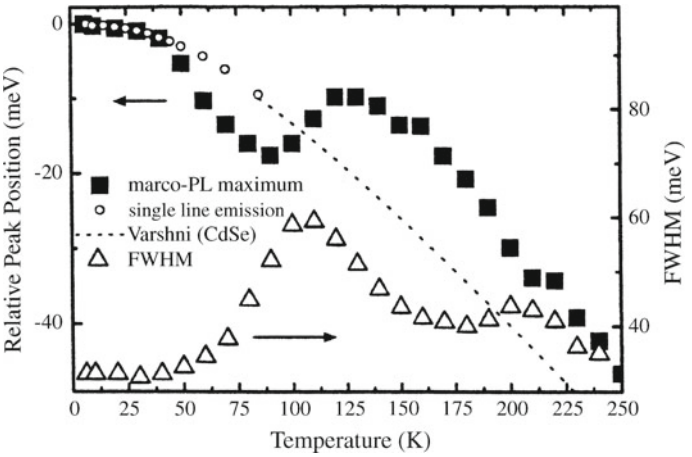


Fig. 23.7 The temperature dependences of the band gap of CdSe according to Varshni's formula, of the narrow emission lines related to individual localization sites (o), of the maximum of the macro-photoluminescence (■) and of the FWHM (Δ) (reprinted with permission from [01W1] ©2001 Elsevier)

Due to the hindered transport within the localized states excitons can "freeze" in the density-of-states tail to form non-thermal distributions. But since broadening by well-width fluctuations in II–VI systems, like in coherently strained narrow quantum wells or in superlattices is sometimes broader than the energy of the optical phonons, and since the coupling of the excitons to the phonons is stronger than in the III–V system AlGaAs, one can observe phonon-assisted relaxation of excitons within these tail states [91C1, 92K1]. Localization can also occur due to the formation of magnetic polarons in dilutes semi-magnetic II–VI structures containing Mn [10Y1]. We will return to these issues later.

If two localization sites of excitons with almost identical energies are spatially so close to each other in one quantum well that they interact slightly, one finds the phenomenon of *level repulsion*. This level repulsion can be observed if high spatial and spectral resolutions are used simultaneously. For theoretical calculations based on random matrix theory and experimental evidence using auto-correlation of SNOM spectra, or time-resolved resonant Rayleigh scattering and speckle analysis see, e.g., [03K3, 03M1, 03N1, 03Z1].

The investigation of the reflection spectra of quantum wells is difficult due to inhomogeneous broadening and due to the fact that a single layer of typically 10 nm in width contributes much less to a resonance signal than a bulk sample. Nevertheless, excitonic reflection spectra can be observed in samples with sufficient quality (see, e.g., [84S1, 84S2, 88U1, 96S1, 01L1]).

In Fig. 23.8 we give an example of a single GaAs well embedded between superlattices as barriers. Under normal incidence, the typical reflectivity is around 30% for the AlGaAs system away from the resonance, and the heavy- and light-hole excitons

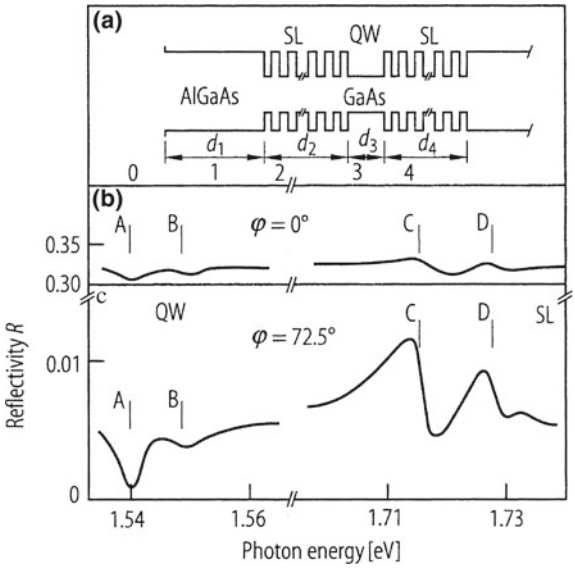


Fig. 23.8 Reflection spectra of a single GaAs quantum well with an $Al_{0.32}Ga_{0.68}As$/GaAs superlattice as barrier (**a**) and the reflection spectra for normal (**b**) and oblique (**c**) incidence (reprinted with permission from [88U1] ©2006 John Wiley and Sons)

in the well (A, B) and in the superlattice (C, D) produce only weak structures. To enhance them, the authors of [88U1] illuminated the sample with linearly polarized light under oblique incidence close to Brewster's angle (Fig. 23.8c). This procedure reduced the background reflectivity and thus enhanced the exciton resonances. Note the different scales in Fig. 23.8b, c on the ordinate.

Exciton polaritons also exist in quantum well structures, however, inhomogeneous broadening often prevents the observation of finer details. In Fig. 23.9 we show calculated in-plane polariton dispersions for the heavy- and light-hole exciton resonances of an ideal single QW without any disorder using the assumption that the background dielectric constant of the well is equal to the dielectric constant of the barrier.

The polariton dispersion in the barrier material is shown by the dotted, almost vertical line. There exists a lower polariton branch that looks similar to the bulk material, with one difference being that the longitudinal-transverse splitting Δ_{LT} depends on the well width l_z and on $k_\parallel$ according to

$$\Delta_{LT} \sim l_z \cdot k_\parallel \tag{23.6}$$

since there is no coupling of light to excitons in a single well with vanishing thickness, nor for infinite wavelength, i.e., $k_\parallel = 0$. The lower polariton branch corresponds to a guided mode since the refractive index below the resonance is higher in the well than in the barrier. There is a longitudinal mode but no upper polariton branch in the usual sense, since the refractive index above the resonance is smaller than that of the barrier. Consequently, there is a radiative or anti-guided mode, which looses its amplitude rapidly (typical in 10 ps) by radiating into the barrier. If the background dielectric constant of the well exceeds that of the barrier, an upper polariton branch

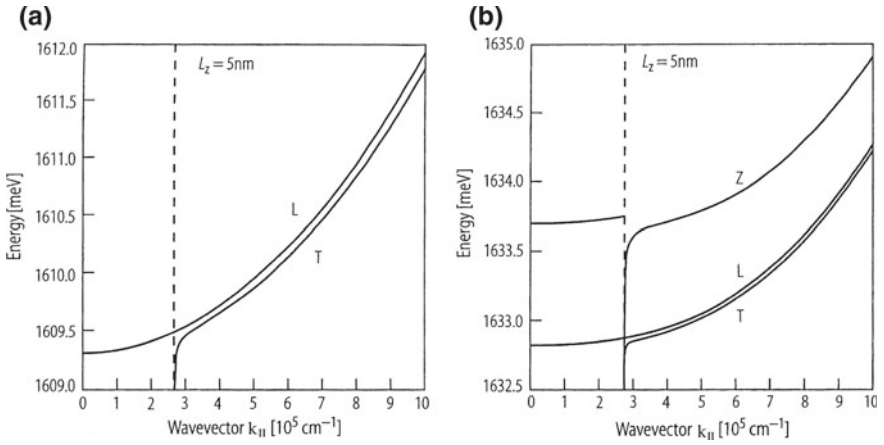


Fig. 23.9 The calculated in-plane polariton dispersion for an ideal 5 nm-wide single GaAs quantum well between $Al_{1-y}Ga_yAs$ barriers for the heavy (**a**) and light (**b**) hole exciton (only dipole allowed states are shown) (reprinted with permission from [93J1] ©1993 American Physical Society https://doi.org/10.1103/PhysRevB.48.1669)

may appear as well [94J1]. The light-hole resonance shows a splitting in the Γ_5 and Γ_4 state. The Γ_5 hh and lh states are dipole allowed for $\boldsymbol{E} \perp z$, where z is the growth direction, and show up in usual transmission spectroscopy, while Γ_4 is dipole allowed for $\boldsymbol{E} \parallel z$ only, and is therefore not seen in the usual geometry $\boldsymbol{k} \parallel z$; $\boldsymbol{E} \perp z$.

From the experimental side, time of flight measurements have been performed, which show a clear decrease of the group velocity around the resonances [90O1], but much less pronounced than for bulk material. Further discussions on excitonic polaritons in quantum wells may be found in [84S1, 84S2, 89K1, 90A1, 90T1, 92T1, 92T2, 95A1, 95B1, 95C1, 98K3, 00M1, 01K2].

For the geometry $\boldsymbol{k} \parallel z$ and $\boldsymbol{E} \perp z$, a stack of quantum wells acts as an ensemble of localized oscillators with infinite translational mass and with a very heavy one in the case of a superlattice. This occurs in the sense that these structures are effective media, as long as the distances between the wells and the well thickness itself are considerably lower than the wavelength of light. For radiative coupling between adjacent quantum wells see, e.g., [90I1, 94C1, 96S2, 96W1, 96W5, 98H1].

The research on polaritons has intensively revived for quantum wells incorporated in micro-cavities. One finds strong coupling of the excitons to cavity modes leading to so-called *cavity polaritons* accompanied by a wealth of coherent phenomena like condensation and superfluidity (see e.g., [03K1, 07D1]). We will discuss these topics in detail in the second volume of this textbook.

We have mentioned the inversion between the magnitudes of the longitudinal transverse splitting Δ_{LT} and the singlet triplet splitting Δ_{st}. For bulk semiconductor samples with direct, dipole-allowed gaps one finds for the 1S Wannier exciton (polariton) $\Delta_{LT} > \Delta_{st}$ while Δ_{LT} decreases with decreasing volume fraction of the material showing the resonance Δ_{st} increases with increasing confinement. This is

already true for (multiple) quantum wells and even more so for quantum dots. See also Fig. 20.3 or [01L1].

Intra-excitonic transitions have been observed, e.g., in GaAs/$Al_{1-y}Ga_yAs$ quantum wells in [89O1, 96C1, 96S3]. For the observation of this transition as a proof for the presence of excitons see [04C1].

In quantum-well structures with type-II band alignment the Coulomb interaction is decreased due to the spatial separation of electrons and holes in the two materials and the wave-function overlap is reduced to the interface region. Excitons in such structures are also said to be indirect in real space. We will discuss the optical properties of such excitons in superlattices in the next section. There we will also address *indirect excitons* in coupled quantum wells where electron and hole are separated by a barrier layer.

Quantum-well structures can also be grown from organic semiconductors [90S1, 03P1, 09M1, 10R1] or combinations of organic and inorganic layers [05T1, 09P1].

23.2 Excitons in Coupled Quantum Wells and Superlattices

In Sect. 17.2 we have introduced symmetric and asymmetric coupled double quantum wells (see Fig. 17.7). Due to the coupling across the thin barrier these double QW show level splitting lifting state degeneracy (see Fig. 23.10), carrier or exciton transfer via tunneling between the wells, as well as direct (electron and hole in same well) and indirect (electron and hole in adjacent wells) excitons.

Excitons in coupled quantum wells (CQW) can thermally populate the excitonic states in both QWs with a common temperature when the wells are coupled [94L1]. This is shown in Fig. 23.11 by luminescence spectra recorded for an asymmetric CQW consisting of two different GaAs wells separated by a thin AlGaAs barrier. The coupling occurs via mixing of the light holes in the wide well with the heavy

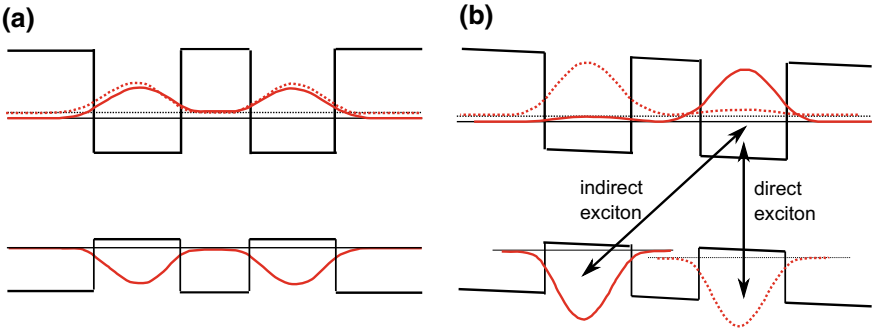


Fig. 23.10 Symmetric coupled quantum well (coupling for the holes is still negligible) (**a**), and the same with with an applied electric field leading to spatially direct and indirect excitons (**b**); for an explicit calculation of the carrier's probability density see, e.g., [90A2]

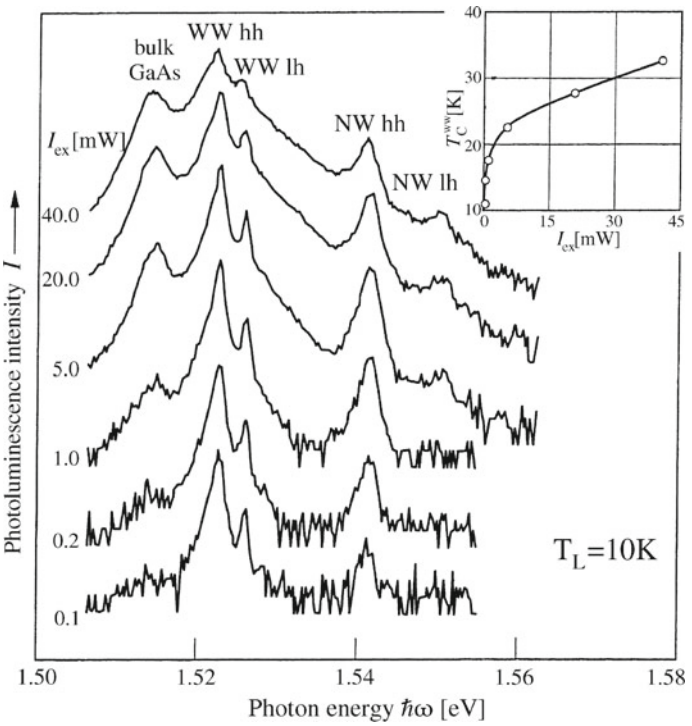


Fig. 23.11 Luminescence spectra of an asymmetric CQW with 9 nm (NW) and 18 nm wide (WW) GaAs wells separated by a 3 nm $Al_{0.32}Ga_{0.68}As$ barrier at $T = 10\,K$ for various cw excitation conditions. The inset shows the temperature of the excitons deduced from the exponential slope of the high-energy luminescence tail (reprinted with permission from [94L1] ©1994 Elsevier)

holes in the narrow well. With increasing pump power, the exciton gas is heated (the effective exciton temperature is given in the inset as a function of pump power) resulting in wider emission bands and increasing importance of the higher states proceeding from the hh exciton of the wide well (WW) to the lh exciton of the narrow well (NW).

Application of an electric field perpendicular to the plane of the QWs can separate electrons and holes in CQW (see Fig. 23.10). Thereby it is possible to switch the nature of the lowest bright exciton states from spatially direct to indirect [88C1]. This transition is shown in Fig. 23.12 [90A2].

Spatially indirect excitons have long lifetime resulting from the reduced overlap of their wave functions, a long diffusion length and a week repulsive potential between the pairs due to their parallel electric dipole moments, and show condensation phenomena [98Z1, 02K1, 05V1, 06G1, 07G1, 17B1]. We will discuss these dynamic and condensation phenomena in the second volume of this text book.

We have already discussed in Sect. 17.2 that within the band-structure models based on the one-electron approximation minibands form along the direction of

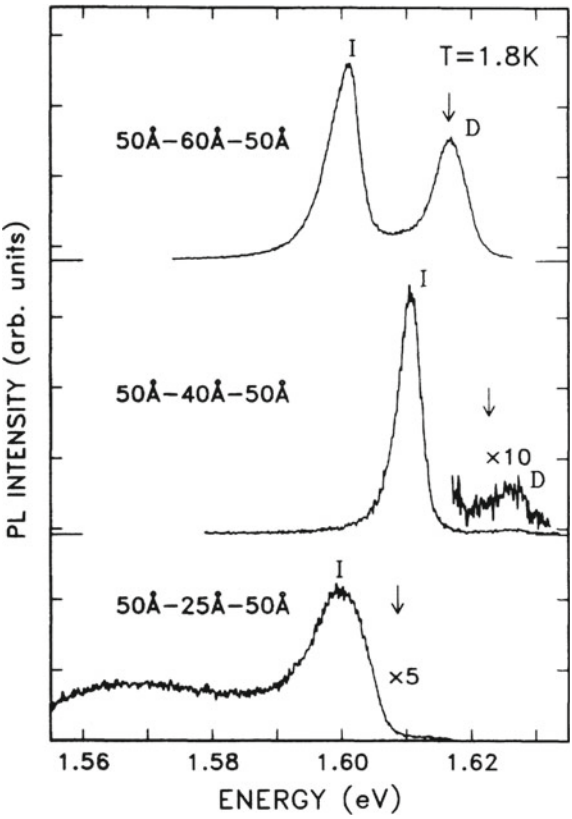


Fig. 23.12 Time-integrated photoluminescence spectra of symmetric GaAs DQWs with different barriers and with applied electric fields. The arrows indicate the peak positions of the luminescence lines under flat-band conditions. The different energy shifts of the indirect luminescence peaks relative to the direct ones reflect the difference in barrier width. The low-energy luminescence band in the lowest spectrum is associated to recombination at impurities (reprinted with permission from [90A2] ©1990 American Physical Society http://doi.org/10.1103/PhysRevB.42.9225)

periodicity of superlattices (SL). The same happens for exciton states in such coupled systems.

The optical properties of type-I SL do not deviate very much from those of MQW [91C2, 91E2] except for some broadening due to the formation of minibands and the increasing importance of monolayer fluctuations with decreasing l_z. In superlattices obviously also fluctuations in the barrier width contribute to disorder. These broadening mechanisms typically mask the details of polariton formation in such structures (see discussion in Sect. 23.1 and [95B1, 97P1]).

The change of optical spectra as a function of layer thickness is demonstrated in Fig. 23.13 for a series of ZnSe/ZnS superlattices. The increasing blue shift with decreasing well width is obvious. Also, the increasing spectral broadening of the light-hole and heavy-hole features are due to the increasing relative importance of well-width and barrier-width fluctuations as well as alloying.

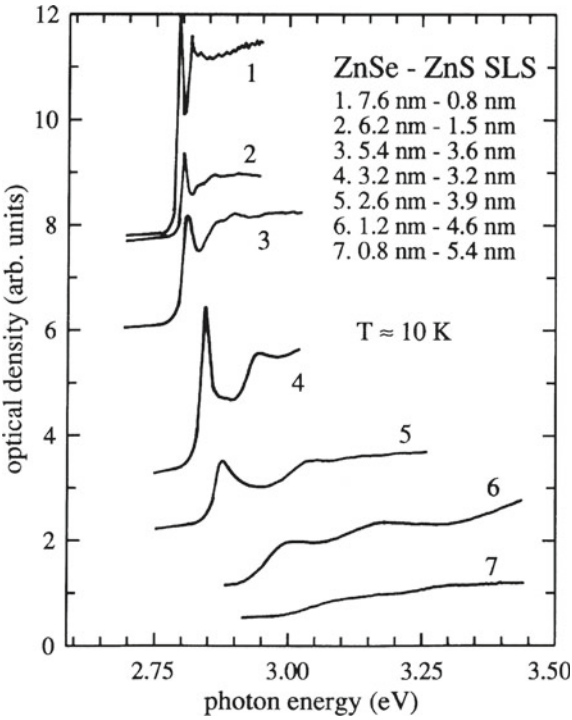


Fig. 23.13 Absorption spectra of a series of ZnSe/ZnS superlattices with various well (ZnSe) and barrier (ZnS) widths (reprinted with permission from [92O1] ©1992 Elsevier)

The CdS/ZnSe system forms superlattices and quantum wells of type-II band alignment with electron confinement in the CdS layers and a remarkably large conduction-band offset by 750 ± 50 meV. The second remarkable finding both in quantum wells and superlattices of this system is the wide tunability of the energy position of the emission peak from about 2.0 to 2.6 eV [99D1, 00D1, 01S3, 04P1].

In the superlattice, the electron and hole wave functions have a stronger overlap than in single and multiple quantum wells of the same CdS well width due to the confinement of the holes in ZnSe in the first case. Additionally, the overlap increases in both cases with decreasing width of the CdS well. Consequences are a more strongly pronounced hh exciton-feature in the superlattice compared to quantum wells, a higher oscillator-strength and a faster luminescence decay time (see [00S1, 04K1]).

In quaternary type-II systems like BeTe/ZnSe or CdS/ZnSe the overlap of electron and hole wavefunctions is limited to the interface region. Depending on the type of interface the bonds can be oriented in (110) or (1–10) directions. Depending on this orientation the luminescence can be polarized preferentially in one of these two directions, although both constituents have cubic symmetry (see, e.g., [00Y1, 04K1]).

The II–VI compounds usually have a rather strong lattice mismatch (see, e.g., Appendix B and [04K1]). As a consequence only rather thin films (a few monolayers) can be grown to avoid the formation of misfit dislocations. The results are so-called *short-period* or *strained-layer superlattices* where the two different materials forming the wells and the barriers adopt a common, intermediate lattice constant.

The strain in such superlattices produces, via the piezo-effect in the layers, electric fields in the growth direction. The piezo-effect requires at least partially ionic binding and the absence of inversion symmetry in the crystalline structure. Both conditions are fulfilled, e.g., for the zincblende- and wurtzite-type III–V and II–VI compounds. Since these built-in fields tilt the electronic bands and shift the exciton energies, such superlattices are also known as *Stark-superlattices*. The piezo-effect is generally stronger in the more ionic II–VI compounds compared to III–V materials and in the wurtzite structure compared to the zincblende one. We present here example data for wurtzite CdSe/CdS superlattices from [94G1, 94G2, 94L2, 94L3, 95L2]. Data for other materials are found, e.g., in [93C2, 95L1, 04K1].

CdSe/CdS has a slight type-II band alignment with a conduction-band offset of 100 ± 100 meV. In Fig. 15.11a we show the conduction and valence band edges including the envelope wave functions of electrons and holes in the first two minibands at $k_z = 0$, the temporal evolution of the emission spectra after ps excitation with 2.46-eV photons, a fluence of 0.3 mJ/cm^2 and the calculated shift of the various transitions with time together with experimental data of the emission maximum.

The conduction-band and valence-band edges are tilted by alternating piezoelectric fields with a field strength of 2×10^8 V/m. Figure 23.14b gives the same results calculated self-consistently for a carrier density per layer of 1.2×10^9 cm^{-2}. Obviously the carrier separation into the different layers produces a strong depolarisation field that almost compensates the piezo-fields under the conditions given in Fig. 23.14b. As can be seen by comparing Fig. 23.14a, b both the transition energies and the electron–hole overlap (i.e., the transition matrix element) depend strongly on the carrier density. This effect can be nicely seen under pulsed (Fig. 23.14c) or cw excitation [94L3].

Both the decay rate and the position of the emission maxima shift to lower values with increasing time after the excitation pulse, leading to an almost logarithmic dependence, as seen in Fig. 23.14d. The temporal evolution of various calculated inter-band transitions are compared to the one of the emission maximum.

In addition, the position of the chemical potential μ of the electron hole pair system is given, i.e., the energetic distance between the quasi-Fermi levels of electrons and holes. At the beginning, it reaches values in the miniband, i.e., there is population inversion. Indeed stimulated emission has been observed from the transition $e2 \rightarrow hh1$, which has a larger oscillator strength than the $e1 \rightarrow hh1$ process [95L2].

We now want to return to III–V SL. Actually the interfaces between the layers in as SL do not need to be flat. GaAs/AlAs SL with a periodic corrugation of the interfaces have been grown on highly indexed substrate surfaces like (113). For some time the aspect of quantum well-wire superlattices has been discussed for this type

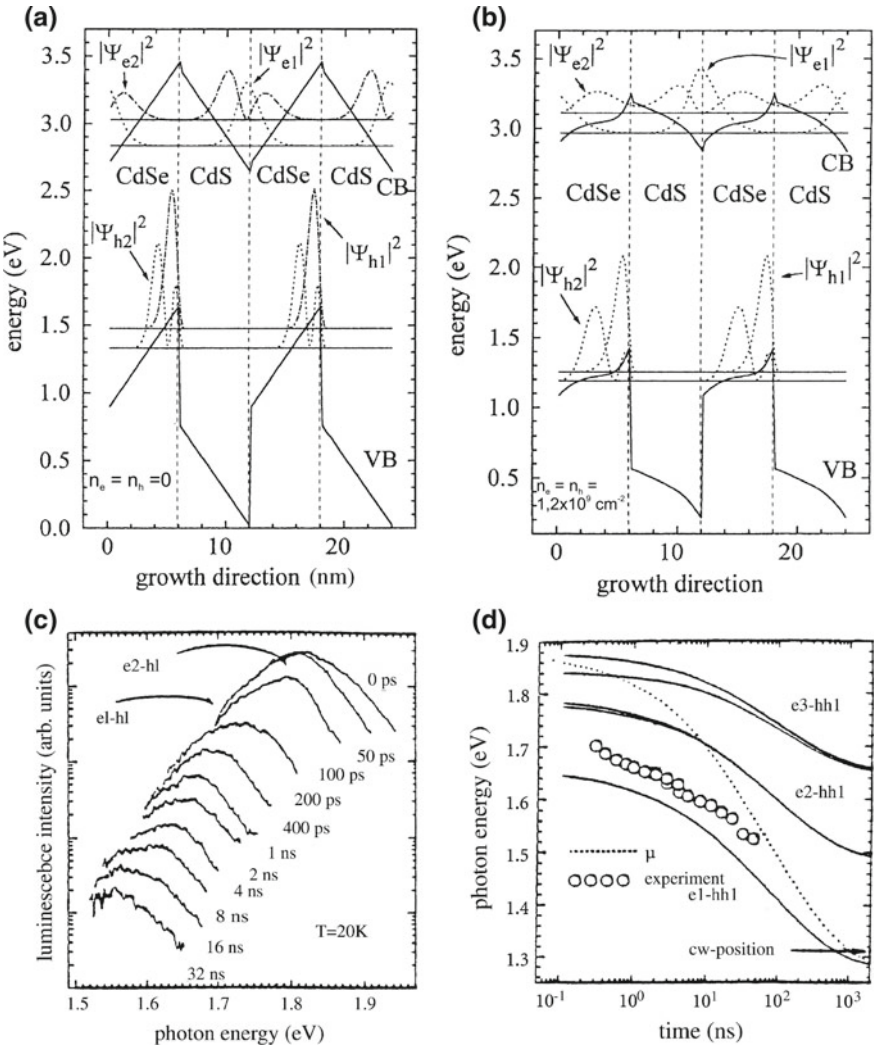


Fig. 23.14 The conduction and valence band edges of CdSe/CdS superlattices influenced by the piezo-field (**a**) for zero carrier density and for a density of $n_e = n_h = 1.2 \times 10^9$ cm^{-2} (**b**) (**a** and **b** reprinted with permission from [94L3] ©1994 Elsevier) the temporal evolution of the luminescence spectra after ps excitation (**c**) (reprinted with permission form [95L2] ©1995 American Physical Society http://doi.org/10.1103/PhysRevB.51.9922) and the experimental and calculated shifts of emission maxima with time (**d**) (according to [95L3])

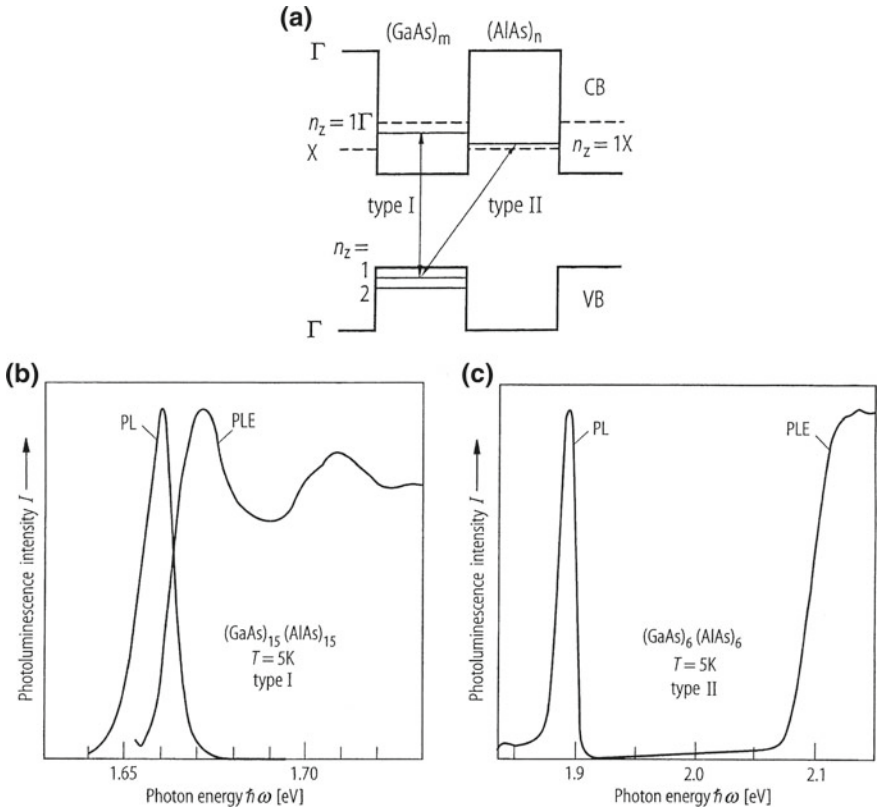


Fig. 23.15 The band alignment of a $(GaAs)_m(AlAs)_n$ short-period superlattice of type II showing the conduction and valence band edges at Γ (——), the conduction-band edge at X (- - - -) and the quantized electron and hole states (**a**), PL and PLE for a type I (**b**) and a type-II superlattice (**c**) (reprinted with permission from [90G1] ©1990 Elsevier)

of structure. However, eventually it has been found that the quantum wire aspect in these structures is marginal or completely absent [96L1].

The next class of superlattices are *short-period superlattices* which consist of few monolayers only of each constituent material. The generally used nomenclature is $(GaAs)_m(AlAs)_n$ where m and n give the number of respective monolayers.

Actually, a transition from type-I to type-II band alignment can be observed in short period GaAs/AlAs SL. GaAs is a direct-gap material and AlAs an indirect one with a larger gap. If the width of the GaAs well is made smaller and smaller, then the $k = 0, n_z = 1$ conduction-band state shifts to higher and higher energies until it is situated above the lowest conduction-band state in AlAs which is situated at the X-point Brillouin zone boundary. This situation is shown schematically in Fig. 23.15.

In contrast to the type-II GaSb/AlAs SL, the short period GaAs/AlAs SL are indirect in both real space and in $\boldsymbol{k}$-space. The peculiarities of absorption in the well, relaxation of the electrons into the "barrier", and subsequent recombination are discussed in detail in [90G1]. Optical properties are shown in Fig. 23.15, where the photoluminescence excitation spectra and the luminescence spectra are given for a type-I and a type-II GaAs/AlAs SL. In the case of a type-II arrangement, the optical matrix element for the transition between the direct maximum in the valence band of GaAs to the indirect minimum in the conduction band of AlAs is so small that the transition hardly shows up in the absorption spectrum. The thermal population factor is responsible for the fact that the low-temperature luminescence nevertheless originates essentially from this transition (Fig. 23.15c).

Experimental Technique: Photoluminescence Excitation Spectroscopy (PLE)

The exciton/carrier relaxation to the radiative states opens up the possibility for a convenient investigation of excitonic (including excited states) and band-to-band transitions in semiconductor nanostructures. Such information can in principle also be gained from the absorption spectrum. In practice however, this typically requires to remove the substrate on which the nanostructure has been epitaxially deposited. Further, the absorption of a nanostructure having a thickness of a few nanometer is quite small and difficult to measure. These problems are overcome by the method of photoluminescence excitation spectroscopy (PLE). With this method one records the PL intensity as a function of the exciting light wavelength [06K1].

The detection wavelength is chosen to be at the low-energy side of the excitonic PL within the tail-state emission. Thus the detection point is situated below the maximum of the close-by excitonic absorption resonance (see arrow in Fig. PLE-1). The wavelength of the excitation source (typically a narrow-linewidth cw laser or a narrow line filtered by a monochromator from a white light source; in both cases the light intensity is kept strictly at a fixed level) is now tuned along the photon dispersion in Fig. PLE-1 across the excitonic resonances and the continuum. Around the intersections of the photon line and the exciton dispersion the excitation light is absorbed. The photo-generated carriers/excitons relax to the radiative states within the PL line. Provided this relaxation is equally efficient for all excess energies one finds that the PL intensity, measured at the detection point, is directly proportional to the absorbance at the energy of the excitation. Thus the PLE spectrum (i.e., the PL intensity as a function of excitation energy) resembles directly the absorption spectrum (red line in Fig. PLE-1) [06K1].

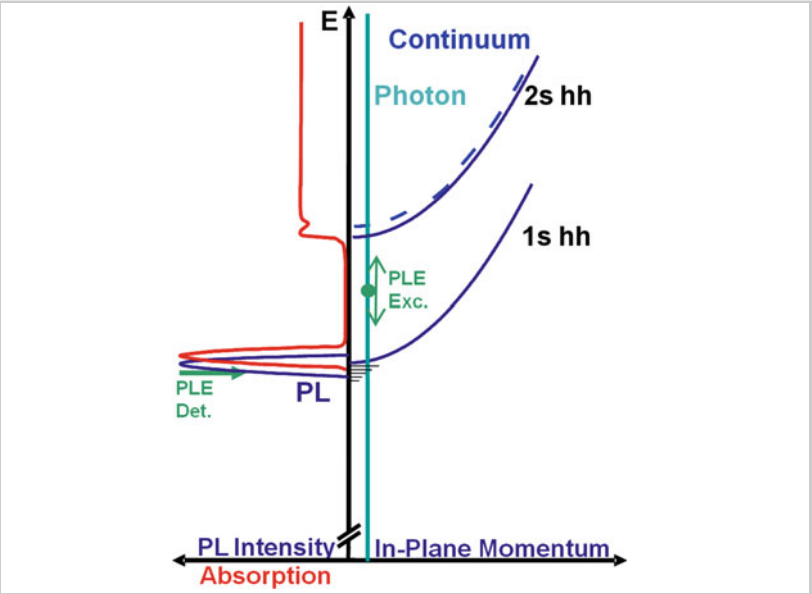


Fig. PLE-1 Dispersion for the lowest exciton states typical for quantum wells of zincblende semiconductors like GaAs and ZnSe (right hand side). The dashed area below the 1s hh parabola indicates the presence of localized states due to disorder. The exciton continuum states are found in the area enclosed by the dashed parabola. Also plotted for comparison is the dispersion of a photon in the weak coupling approximation. The left hand side shows schematically the photo-luminescence and absorption spectra related to these exciton states (adapted with permission from [06K1] ©2006 Springer Nature)

Since relaxation of the carriers/excitons can also happen through alternative channels (e.g., non-radiative), it is typically not possible to deduce the value of the absorption coefficient. But the energy position of optical transitions can be reliably determined.

Examples of PLE spectra are given, e.g., in Figs. 23.5 and 23.15.

Large efforts have been made to grow strained Ge/Si SL. There are considerable technical problems in growing good quality samples due to the different lattice constants. The idea is that a similar zone-folding mechanism as outlined in connection with phonons (see Sect. 13.3) should yield a direct electronic bandstructure. This would allow the highly developed Si technology to be used also for light-emitting devices. Although direct-gap photoluminescence has been achieved in indirect-gap samples suitably high luminescence yields have not been achieved (for a data collection see part C3 of [01L1]).

To conclude this section we briefly return to the concept of nipi structures or doping superlattices introduced in Sect. 17.2. The probability of transitions between electron states in the conduction-band minima and holes in the valence-band maxima

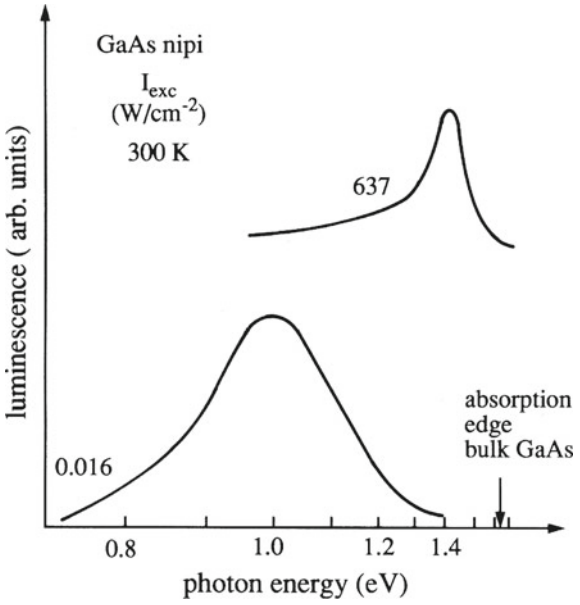


Fig. 23.16 Luminescence spectra of a nipi structure for two different pump powers compared to the RT absorption edge of bulk GaAs (adapted with permission from [86D2] ©1986 Elsevier; also compare with Fig. 23.14)

is low because of the spatially indirect bandstructure. Consequently strong absorption sets in only around the energy of the direct gap, i.e., at an energy corresponding to that between the bands at the same place, modified by the Coulomb interaction. The carriers are spatially separated after their creation resulting in a very long lifetime. The radiative recombination which can nevertheless occur due to the small overlap between electron and hole wavefunctions in their respective spatial band minima is strongly Stokes-shifted with respect to the spatially direct absorption edge. We give an example in Fig. 23.16, which shows the emission spectra for two different pump intensities and the position of the absorption edge of bulk GaAs. An analogous shift of the emission with excitation is also seen in piezo-superlattices see Fig. 23.14 or [93C2, 94L3, 95L2]. The non-linear electro-optical properties of nipi structures are interesting for device applications [86D1, 92M1, 93L1].

23.3 Excitons in Monolayer Semiconductors

In Sect. 17.3 we have described the bandstructure of monolayer (ML) semiconductors like transition-metal dichalcogenides (TMDs). They have a triangular lattice and a hexagonal Brilloiun zone with direct gaps situated at the K and K′ points, i.e., the zone edges. Due to spin-orbit interaction there is a splitting of both conduction and valence bands and a locking of spin and *valley degrees of freedom* (see Fig. 17.12).

The optical transitions between the various valence and conduction bands are actually excitonic transitions. Excitons in TMDs are strongly bound due to large

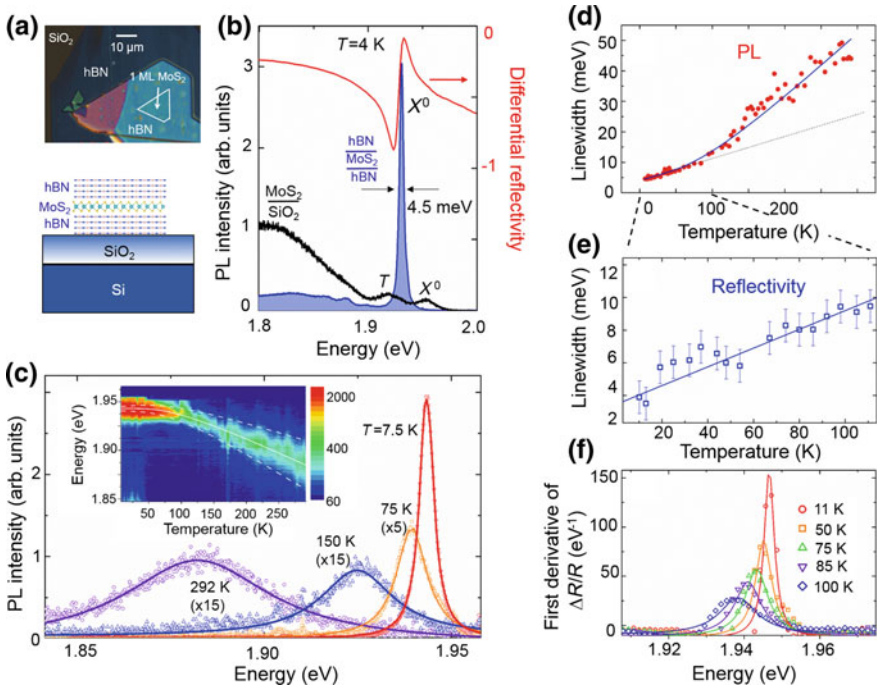


Fig. 23.17 Top: Optical microscope image of van der Waals heterostructure hBN/ML MoS_2/hBN. Bottom: Schematic of sample (**a**); cw-PL spectrum (filled curve) and differential reflectivity spectrum (red curve) at $T = 4$ K for a capped ML. Also shown is the PL spectrum of an uncapped MoS_2 ML deposited directly onto the SiO_2 substrate (black curve) measured under the same conditions (**b**); PL spectrum of a capped sample for selected sample temperatures. The inset shows a color map revealing the temperature evolution of the spectrum's intensity. The white full line is fit with a hyperbolic cotangent relation to the evolution of the peak position, whereas the dashed lines are a guide to the eye indicating the linewidth (FWHM) (**c**); Temperature evolution of the exciton linewidth extracted from the PL spectra. The solid line is a fit according to equation given in Fig. 21.13, and the dashed line represents the linear term which dominates at low temperatures (**d**); Linewidth extracted from the first derivative of the differential reflectivity as a function of temperature. The linear fit is consistent with a broadening induced by scattering with low-energy acoustic phonons (**e**). The first derivative of the differential reflectivity for selected temperatures, centered around the exciton absorption (**f**) (republished under Creative Commons Attribution 4.0 International license from [17C1])

effective masses of electrons and holes, the reduced screening described above for (quasi-)2D systems, and carrier confinement in the single monolayer [15W1, 16M1]. The excitonic binding energies are in the range of 400 to 700 meV and the excitonic Bohr radius is about 1 nm so that the excitons are robust species at room temperature. For a table of exciton binding energies and band parameters see [16M1], a theoretical treatment of exciton binding in effective mass approximation is given in [13B1].

Still, excitonic transitions of monolayer TMDs exfoliated on SiO_2 show typically significant inhomogeneous broadening of the order of 10 meV (50 meV) at

$T = 4$ K (300 K). The broadening can be efficiently reduced by encapsulating the ML by layers of hBN [17C1, 18S1]. This allows (similar to the case of quantum wells) the access to intrinsic excitonic properties as shown in Fig. 23.17 for the case of MoS_2. The exciton energy position follows the typical behavior found in semiconductors. The linewidth of the exciton in photoluminescence (PL) and derived from differential reflectivity shows the expected linear dependence on temperature due to interaction with acoustic phonons. The coupling to optical phonons dominates at higher temperatures. The close energy distance between bright and dark exciton states might be responsible for the non-monotonous behavior of the PL intensity with temperature [17C1].

The nature of the fundamental excitonic transition in TMDs is strongly influenced by the intra-valley exchange interactions. The bright excitons are lowest in energy in $MoSe_2$ and $MoTe_2$ while the dark exciton state is lowest in WSe_2 and WS_2. The dark-bright splitting is close to zero in MoS_2. For a theoretical treatment and discussion see [16E1]. Dark excitons can be visualized in optical spectroscopy by coupling to surface plasmons when a van der Waals heterostructure is places on a metal substrate [17Z1, 18P1].

The strong spin-orbit interaction leads in optical spectra to the appearance of two dominant excitonic lines called A and B excitons. The separation of the two lines is due to the valence-band spin splitting (for the band-structure details at the K points see Fig. 17.12. A feature called C exciton observed in WS_2 is related to higher gaps between band extrema close to the Γ-point [15Z1]. For the signatures of the A and B excitons in reflectivity spectra and in the spectra of the dielectric function (ϵ_1 and ϵ_2) calculated from Kramers–Kronig analysis see [14L1].

The exciton in TMDs are of Wannier-type but their series of excited states actually deviates from the one expected from a pure 2D hydrogen model [14C1, 14H1, 14Y1, 15Z1]. The reason is the so-called *dielectric confinement*, i.e., a distance-dependent strength of the screening which results in stronger binding for larger electron–hole separation compared to the one expected from the 2D Rydberg series [14Y1]. Theoretical absorption spectra for the A- and B-exciton series in WS_2 which are in agreement with two-photon absorption data are shown along with exciton wave-functions in real space in Fig. 23.18.

The exciton orbital functions in Fig. 23.18 still resemble the ones of the 2D hydrogen atom. However, the deviations from the hydrogen series become obvious from the order of the s, p, and d states for the same main quantum number. A consequence of this deviation is further, that a simple determination of the exciton binding energy from the separation between excitonic transitions is not possible [15Z1]. The states with orbital quantum numbers beyond s are of course optically dark and only accessible in multi-photon absorption.

For the evolution of excitonic photoluminescence spectra as a function of number of monolayers see [13T1]. The intensity of the PL spectra reflects the transfer to indirect bandgaps with increasing thickness.

In the presence of additional free electrons excitons can bind an electron to form negative trions [12M1] with large binding energies between 20 and 30 meV. For a theoretical treatment of trions in TMDs see [13B1]. There is doubt about this

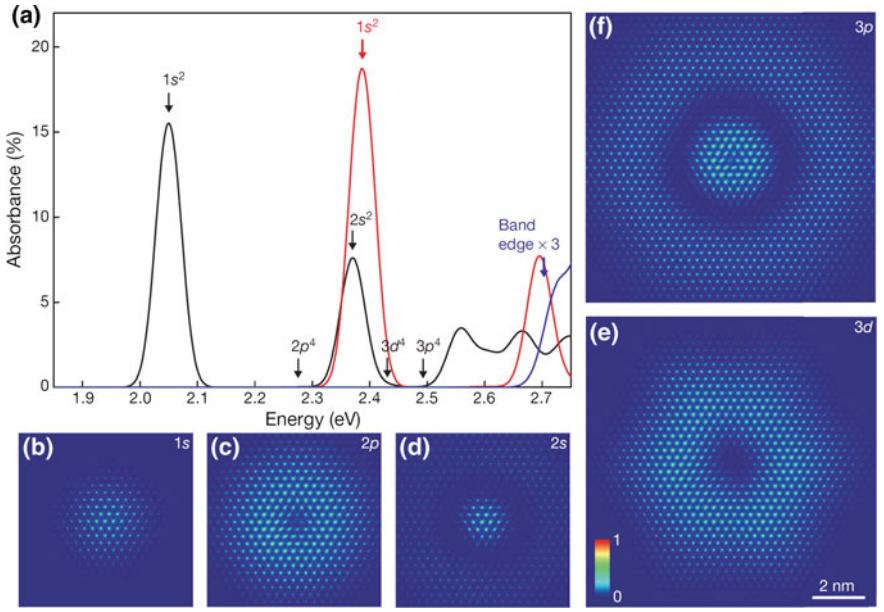


Fig. 23.18 Optical absorption of the A (black) and B (red) exciton series including electron–hole interaction and without (blue). The quasi-particle bandgap is indicated by the blue arrow, the excitonic states by black and red arrows, respectively. The orbital notation of a 2D hydrogen atom is adopted to label the exciton states, the degeneracy labels in the superscript include both the degeneracy of valleys and orbital angular momentum (**a**). Modulus squared plots of the real-space exciton wave-function projected onto the WS2 plane. The color scale is the normalized wave-function probability (**b–f**) (reprinted with permission from [14Y1] ©2014 Springer Nature)

interpretation since the trions are observed only in the presence of a large density of free carriers. An alternative ansatz interprets the two peaks in the spectra as exciton-polaron branches resulting from a dressing of the exciton by interaction with the electron Fermi sea [17E1].

The formation of polaritons is well studies in monolayer semiconductors in particular for plasmon and phonon polaritons (see [16B1, 17L1] for reviews). Plasmon polaritons are here considered to be in the extreme limit of short-range surface plasmon polaritons, which propagate confined by two surfaces in thin metallic films. Here, the monolayer material like graphene allows wide tunability of the plasmon parameters. Phonon polaritons can show a hyperbolic dispersion like in hBN. This results from the strong birefringence associated with the structural anisotropy which features a sign reversal of the permittivity for in- and out-of-plane directions.

Exciton polaritons have mainly been observed for monolayer materials incorporated in micro-cavities. The spin-valley locking can be preserved leading to *valley polaritons* up to room temperature [17C2, 17D1, 17S1]. The topic of strong coupling of excitons and cavity photons will be detailed in the second volume of this text book.

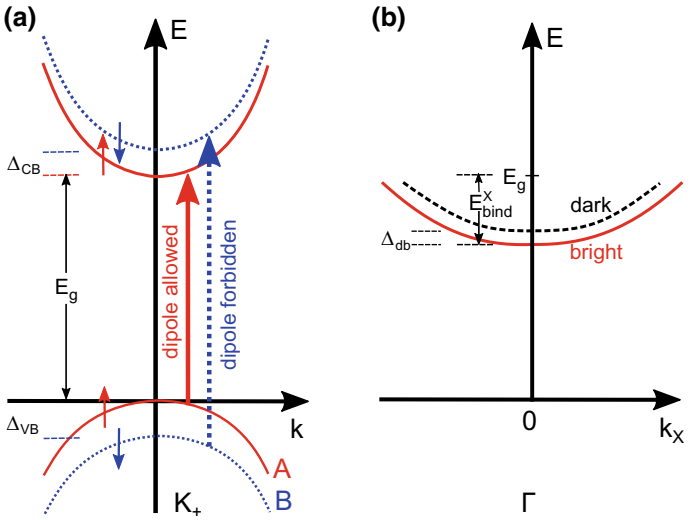


Fig. 23.19 Bandstructure at the K_+ point of a monolayer of MoX_2 in the one-electron approximation; spin orientation of the electrons in the conduction band (CB) and the missing electrons in the A and B valence bands (VB) are given by arrows; E_g is the energy of the fundamental bandgap; Δ_{CB} and Δ_{VB} are the splittings of the respective bands (**a**). Energy dispersion E of the direct bright and dark excitons in MoX_2 as a function of exciton center-of-mass momentum k_X; exciton binding energy E_{bind}^X and the dark-bright splitting Δ_{db} are indicated

Spatially indirect excitons can be excited when layers with with different work functions are combined like MoS_2/WSe_2 or $MoSe_2/WSe_2$. These excitons have long lifetimes. The exciton binding energy can be adjusted as a function of layer distance [16N1]. For the band offsets between various TMDs see [13K1].

Monolayer semiconductors allow studies of 2D excitons in a more well defined way than typically possible in the quasi-2D quantum well structures discussed in Sect. 23.1. Many of the exciting excitonic properties are and will be published in high-impact journals. This does not prevent authors, reviewers and editors to publish, e.g., schematic depictions of excitons (plotting erroneously the constituting electron and hole in their single-particle bands) which are misleading and neglect the nature of the exciton of being a composite particle with its own dispersion given by the center-of-mass motion. A discussion of selection rules for excitonic transitions based only on the bandstructure in the one-electron approximation gets particularly awkward when talking about dark excitons.

We want to close this section by summarizing the properties of free-carrier and exciton states in Fig. 23.19 following the discussions in [15D1, 16E1]. The free carrier bandstructure is given in (a) at the K_+ points in the one-electron approximation introduced in Sect. 14.2 (the corresponding bandstructure at the K_- point follows from time reversal). For dipole-allowed (forbidden) transitions between orbital states the electron spins are oriented parallel (anti-parallel). Since the splitting of the valence

band Δ_{VB} by spin-orbit coupling is much larger than one of the conduction band Δ_{CB}, the Fig. 23.19b depicts only the bright and dark A exciton dispersions. Although electron and hole reside at the K point their difference in momentum $\boldsymbol{k}_\mathbf{h} - \boldsymbol{k}_\mathbf{e}$ is zero so that the parabolic dispersion of the excitons is centered at the Γ point. The ordering of the bright and dark excitons and the magnitude of the dark-bright splitting Δ_{db} are not only given by Δ_{CB} but also by the short-range electron–hole exchange term. Due to magnitude and relative sign of these contributions the ordering is reversed in WX_2 monolayers (see [16E1]).

23.4 Excitons in Quantum Wires and Nanorods

In strictly one- and zero-dimensional cases the binding energy for the exciton diverges. So it is not possible to give general formulas like (23.2) for these situations. One is always limited to numerical calculations which have to explicitly include the finite dimensions of the quantum wire or quantum dot. A variational approach to calculate the exciton binding energy in quantum wires is the use of fractional dimensions. For shallow confinement and wires with large cross-section one can introduce effective wire dimensions. For a more complete description including also narrow wires a microscopic treatment of the consequences of restricted exciton motion is necessary [93C1]. For an analytical theory to treat the Coulomb problem in quasi-2D systems see [03C1].

In Sect. 17.4 we have introduced various possibilities to realize quantum wires, e.g., by laterally structuring quantum wells, by growing in V-groves or on ridges, by cleaved-edge overgrowth producing T-shaped structures or by growing free standing, whisker-like needles, also known as nanorods. Driving force for fabricating this type of low-dimensional structure are the large, spectrally narrow density of states and the wave-guiding properties which are beneficial for lasing applications. Indeed, lasing and even excitonic lasing has been found in quasi-1D structures (see, e.g., [89K2, 93W1, 01H1, 13M1].

In the following, we present a few examples of optical spectroscopy of excitons in quantum-wire and nanorod structures.

The problem encountered in lateral structuring is that the quantization length $l_z \approx 10\,\text{nm}$ in the growth direction of the QW is usually much smaller than the length $l_y \approx 50-100\,\text{nm}$ of the subsequent lateral structuring. As a consequence, the confinement energies resulting from confinement in the two directions are very different. Usually one observes only a narrow modulation of the luminescence or photoluminescence excitation spectra due to quantization in the y-direction. We provide examples for the confinement in the second direction in Figs. 23.20 and 23.21.

Figure 23.20 shows PLE spectra of a 8 nm-wide GaAs quantum well between $Al_{1-y}Ga_yAs$ barriers in the upper trace, showing the well-known heavy and light hole excitons. In the PLE spectrum of the quantum wires these two resonances are slightly blue shifted according to the lateral confinement and for the heavy-hole exciton even the first two laterally confined states could be resolved, labeled hh_{11} and hh_{12} [95S1].

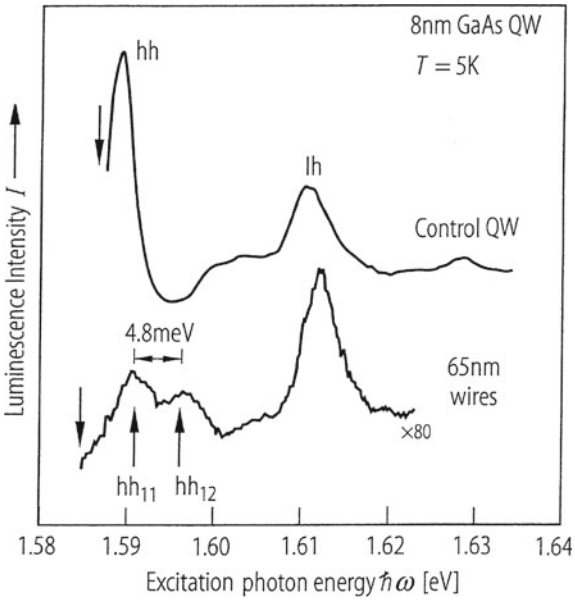


Fig. 23.20 PLE spectra of a 8 nm-wide GaAs/$Al_{1-y}Ga_yAs$ multiple quantum well sample (upper trace) and of 65 nm-wide quantum wires patterned from this sample (lower trace). The PLE spectra have been taken at the spectral positions indicated by the vertical arrows (reprinted with permission from [95S1] ©1995 IOP Publishing https://doi.org/10.1088/0268-1242/10/10/015

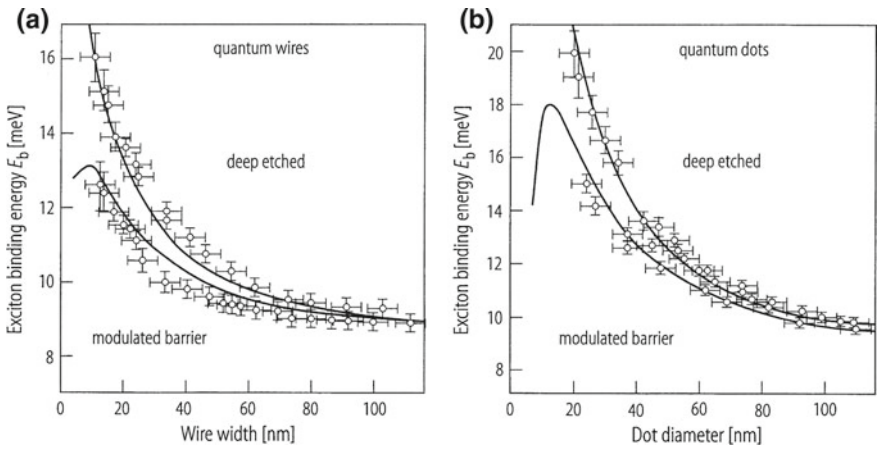


Fig. 23.21 Exciton binding energies as a function of wire width (**a**) or of dot diameter (**b**) for InAs samples (reprinted with permission from [98B1] ©1998 American Physical Society http://doi.org/10.1103/PhysRevB.57.6584)

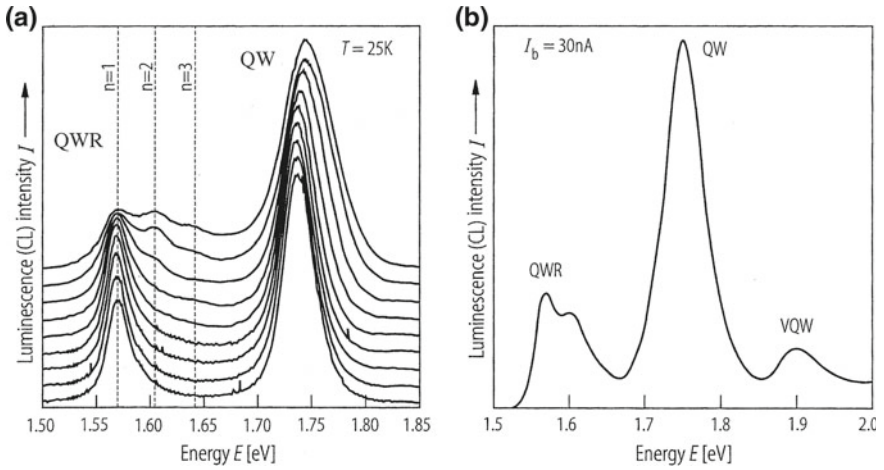


Fig. 23.22 Normalized cathode-luminescence spectra of a single GaAs V-groove quantum wire for increasing electron beam density on the sample at an acceleration voltage of 12.5 kV (**a**) and an overview over the whole spectrum (**b**) (reprinted with permission from [95G1] 1995 ©AIP Publishing https://doi.org/10.1063/1.114923)

Progress in nanolithography is obvious when comparing the data in Fig. 23.21 for InAs wells. Samples have been prepared by either etching through the upper barrier and the well into the substrate (deep etched), or only the upper barrier has been thinned down close to the well (modulated barrier). In both cases, a significant increase of the exciton binding energy from the value of the well at around 8 meV by a factor of 1.5 or 2 has been obtained by additional lateral confinement.

In Fig. 23.22 we show the cathodo-luminescence spectra of a single GaAs V-groove wire. The overview spectrum shows, at the lower energy side, the emission from the quantum wire (QWR) at the bottom of the V-groove. At the higher energy side occurs the emission from the quantum wells (QW) at the side flanks of V-groove. The latter are considerably narrower, and therefore they form the lateral confining potential of the wire-states though they have higher quasi dimensionality $d = 2$. The emission labeled VQW finally results from a vertical quantum well which forms unintentionally in the middle of the V-groove (compare Appendix D).

The lateral confining potential of the V-groove is not as abrupt as, e.g., quantum wires formed by nanolithography, but rather smooth. Therefore the higher states, which are populated at higher excitation current and are seen in luminescence, are in first approximation as equally spaced as in a harmonic oscillator. They are labeled in Fig. 23.21a by $n = 1, 2, 3$.

As a last example we show in Fig. 23.23 the luminescence spectra of two different T-shaped quantum wires formed by cleaved edge overgrowth. Their emission is again red shifted with respect to the emission from the quantum wells on the (001) and

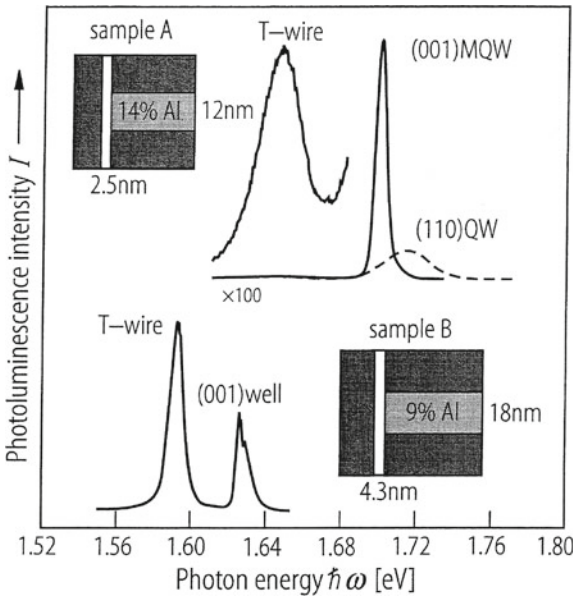


Fig. 23.23 Low-temperature photoluminescence spectra of two different T-shaped quantum wires (reprinted with permission from [96L3] ©1996 American Physical Society http://doi.org/10.1103/PhysRevB.54.14595)

(110) surfaces. Further example may be found, e.g., in [98A1, 01S2]. Exciton binding energies in GaAs/AlGaAs T-shaped quantum wires have a theoretical upper limit of 25 meV for optimized geometry, but experimentally realized values are typically much smaller [01S2].

The anisotropic confining potential for excitons and in particular the anisotropy of the electronic bandstructure in quantum wires is reflected in an anisotropy of optical spectra like PL and PLE. The emission depends on the linear polarization of the excitation light with respect to the wire orientation. For examples see, e.g., [97V1, 98A1, 01H2, 01C1].

Charged excitons (negatively and positively charged trions X^- and X^+) are only weakly localized in quantum wires. Due to the recoil of the remaining charge after recombinations the trions have an enhanced binding energy compared to excitons confined to quantum dots with the binding energy being stronger for the X^- [05O1].

Good access to excitonic properties in quantum wires is given in colloidal CdSe nano-wires. They can be fabricated with narrow wire width. Core/shell structures with CdSe being clad by the large band-gap material ZnS show high luminescence yield. A fine-structure splitting of the exciton due to exchange interaction in the anisotropic wire and a crossing of dark and bright exciton states for small wire diameters is found [05L1].

We finally want to address so-called nanorods. Especially for wide-gap semiconductors like ZnO or GaN nanorods do generally not provide a lateral confining potential narrow enough to produce an appreciable confinement or blue shift of the exciton states. However, these nanorods form efficient wave guides along the axis of the rod for photon-like polaritons (see [06H1, 08F1, 10K1]).

Strong coupling between excitons and whispering-gallery modes in ZnO nanorods of hexagonal cross-section leads to cavity polaritons [08S1]. We will return to the properties of these polaritons in the second volume of this textbook.

23.5 Excitons in Quantum Dots

23.5.1 Excitonic States in Quasi-0D Structures

In the following, we want to review the excitonic properties of quasi-0D structures. We already mentioned in the previous section that the binding energy for the exciton diverges in strictly one- and zero-dimensional cases. The treatment of electronic and excitonic states in quantum dots requires numerical methods which consider the explicit size, shape, surface and environment of the quasi-0D structure. For elaborate discussions of excitons in quasi-0D structures see, e.g., [97W1, 98G1, 98J1, 05B1].

We want to introduce here some simplified basic ideas to deal with the confinement of excitons in (approximately) spherical quantum dots. Such dots actually occur frequently in the cases of semiconductor nano-crystals in glass matrices or of colloidal quantum dots. The following formulae for the quantization energies assume besides spherical shape also simple isotropic and parabolic conduction and valence bands as well as infinite barriers. They further neglect image charges caused by the different dielectric properties of dot and surrounding material [05B1].

One distinguishes for essentially spherical QDs three *quantization regimes*. Regime *I* (*weak confinement*) is obtained for dots with radius $\bar{R}$ larger than the excitonic Bohr radius a_B:

$$\text{Regime I (weak confinement):}\ \bar{R} \gtrsim a_\mathrm{B}\,; \quad \Delta E_\mathrm{Q} < \mathrm{Ry}^*\,. \tag{23.7a}$$

In this case, only the center of mass motion of the exciton, which is described in (20.1a) and (20.4a) by the term $\mathrm{e}^{\mathrm{i}\boldsymbol{K}\cdot\boldsymbol{R}}$, is quantized, while the relative motion of electron and hole given by the envelope function $\phi_{n_\mathrm{B},l,m}(\boldsymbol{r}_\mathrm{e}-\boldsymbol{r}_\mathrm{h})$ are hardly affected. The quantization energy ΔE_Q is given, in the above described approximations by

$$\Delta E_\mathrm{Q}^\mathrm{I} = \frac{\hbar^2\pi^2}{2M\bar{R}^2}\,, \tag{23.7b}$$

where $M = m_\mathrm{e} + m_\mathrm{h}$ is the translational mass of the exciton. This situation is found, e.g., for the Cu halides where a_B is small, or for CdSe QDs with $R \geq 10\,\mathrm{nm}$.

The second regime (*intermediate confinement*) is given by

$$\text{Regime II (intermediate confinement):}\ a_\mathrm{B}^\mathrm{e} > \bar{R} > a_\mathrm{B}^\mathrm{h}\,; \quad \Delta E_\mathrm{Q} \approx \mathrm{Ry}^*\,, \tag{23.8a}$$
$$\text{with}\quad a_\mathrm{B}^\mathrm{e,h} = a_\mathrm{B}^\mathrm{H}\varepsilon\frac{m_0}{m_\mathrm{e,h}}\,,$$

where $a_B^{e,h}$ are the respective radii of the electron and hole orbits in the exciton around their common center of mass. In this case $\bar{R}$ has a value between the radii of the electron orbit and the hole orbit. Thus the motion of the electron is quantized and the hole moves in the potential of the QD and of the space-charge of the electron for the usual situation $m_e \ll m_h$.

Continuing to use the above-mentioned approximations, and neglecting the Coulomb energy between electron and hole, one finds

$$\Delta E_Q^{II} = C\mathrm{Ry}^* \left(\frac{a_B \pi}{\bar{R}}\right)^2, \quad \text{with} \quad C = 0.67\,. \tag{23.8b}$$

The neglection of the Coulomb energy (i.e. of the exciton binding energy) is a rather severe approximation in this regime, because the exciton binding energy increases with decreasing dot radius [88K4]. It is in regime II still comparable to the quantization energy. This situation is actually often realized for QDs of II–VI semiconductors.

The third regime (*strong confinement*) is defined by:

$$\text{Regime III (strong confinement): } \bar{R} \ll a_B^{e,h}; \quad \Delta E_Q > \mathrm{Ry}^*\,. \tag{23.9a}$$

Since the quantization energy scales with $\bar{R}^{-2}$ and the Coulomb energy with $\bar{R}^{-1}$, one can expect as a first guess for this regime

$$\frac{\hbar^2 \pi^2}{2\mu \bar{R}^2} = \Delta E_Q^{III} \gg \Delta E_{Coulomb}^{III} \tag{23.9b}$$

where μ is the reduced mass of electron and hole. The influence of the Coulomb energy can be partly incorporated, e.g., by using [97M1]

$$\Delta E(R) = \frac{\hbar^2 \pi^2}{2\mu \bar{R}^2} - \frac{1.78 e^2}{\varepsilon \bar{R}} + 0.752 E_{Ry}\,. \tag{23.9c}$$

Slightly different formulae and other approaches can be found, e.g., in [86S1, 88K1, 91E1, 01L1, 10H1].

The Coulomb energy increases roughly with $\bar{R}^{-1}$ and the quantization energy with $\bar{R}^{-2}$, so that for sufficiently small values of $\bar{R}$ one should reach a situation where the Coulomb term can be neglected. But the required values of $\bar{R}$ where (23.9b) and (23.9c) hold are actually comparable to lattice constant. In this situation, the whole effective mass concept starts to break down and a description of the dot in terms of molecular orbitals becomes more appropriate. An exhaustive list of references for calculations using the full valence band structure of cubic and hexagonal quantum dots and of calculations for very small clusters where tight binding methods are more appropriate is given in [01L1]. The increase of the electron–hole exchange energy with decreasing dot radius has already been addressed in Sect. 20.3 with Fig. 20.3.

23.5.2 *Spectroscopy of Optical Excitations in Quasi-0D Structures*

Optical transitions in quantum dots follow elaborate selection rules [96W2]. In the simplest picture of dipole-allowed transitions between electrons and holes in an infinitely high spherical potential the lowest transitions would read: $1s_h \rightarrow 1s_e$, $1p_h \rightarrow 1p_e$, ... with identical wavefunctions for electrons and holes and simple expressions for the oscillator strengths. For more a realistic description one needs to account for the following: Coulomb interaction between electron and hole which confines the heavier particle (typically the hole) closer to the dot's center affecting the penetration through the finite barrier and the overlap between electron and hole wavefunctions; bandstructure details of the dot constituent (ZnSe, CdSe, ...) like multiple valence bands, anisotropic band dispersions, band mixing etc. In this situation the fundamental optically allowed transition is essentially unchanged ($1S_{3/2} \rightarrow 1S_{1/2}$, where the subscript is the total angular momentum) but the spectral lines shift and optically forbidden transitions gain oscillator strength.

From the considerations of the density of states (see Fig. 17.1 one would expect QD to have an absorption spectrum consisting of a series of δ-functions or of Lorentzians, if some homogeneous broadening is included. Such behavior is not observed in reality for ensembles of dots since various processes contribute to a broadening of the resonances, e.g., the size distribution of the crystallites considered in Fig. 23.24, the coupling to optical phonons, which may be enhanced in the quantum dots as

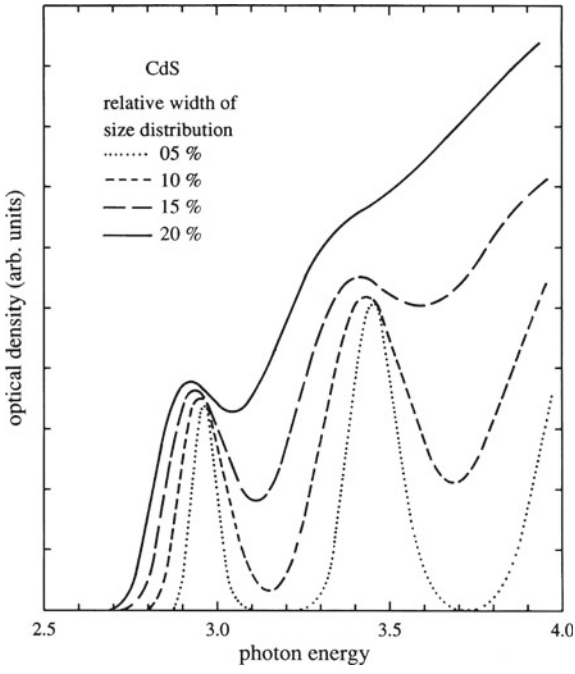


Fig. 23.24 Calculated spectra of the optical density of quantum dots for various relative widths of there size distribution [90K1]

compared to the bulk material, and the influence of impurity, surface, or interface states; see [91E1, 93G1, 93W2, 95W1, 96W1].

The range of the size distribution for nano-crystallites in glass matrices is typically in the range of 10–30%. For colloidal QDs in organic solvents it can be reduced by size selective precipitation to a few percent [95W1]. Quantum dots in MBE-grown layers usually show strong size and composition fluctuations. Ways out of this problem of resulting inhomogeneous broadening are single dot spectroscopy and site- or size-selective spectroscopy, including spectral hole burning. We will discuss these methods in Sect. 23.5.2 and in the second volume of this book.

In Fig. 23.24 we present calculated absorption spectra, with a simplified energy level scheme, showing the influence of inhomogeneous broadening. In Figs. 23.25 and 23.26 we show experimental data taken on quantum-dot ensembles.

Confinement regime I is usually realized for QD made from the copper halides CuCl, CuBr and CuI since the values of a_B for these materials are in the range of 1 nm. Figure 23.25a shows an example for CuBr which clearly illustrates the blue-shift of the $n_B = 1$ exciton resonances from the two uppermost valence bands with decreasing $\bar{R}$.

In Fig. 23.25b we show the blue-shift of the absorption edge of glasses containing CdSe QDs with decreasing values of $\bar{R}$. The calculated transition energies and their oscillator strengths are given by the positions and heights of the vertical lines. For more data see also [92K2].

The luminescence is usually considerably Stokes-shifted with respect to the absorption, in the example of CdSe QDs shown in Fig. 23.26 by roughly 200 eV. This shift can be partly attributed to a relaxation of the carriers into defect or surface states, and partly to stronger phonon coupling resulting in Huang–Rhys factors S^{QD} in the QD of around 1, which is here by about a factor of 5 larger than for the corresponding bulk material. The values of the Huang–Rhys factor are of course strongly dependent on the individual properties of the dots including capping or surface states and have values between 0.2 and 1. For a data collection see [01L1].

The exciton or electron–hole pair ground state usually shows a fine-structure, due to the exchange splitting. The lower state is the dark triplet state, whose optical recombination is spin flip forbidden. This ordering of the states already exists in bulk materials, with the difference that in bulk materials the splitting between the triplet and the transverse dipole allowed singlet states of Wannier excitons is much smaller than the transverse-longitudinal splitting of the singlet. In quantum dots the situation is inverted due to a strong enhancement of the (short-range) exchange interaction caused by three-dimensional confinement, while the oscillator strength of the QD ensemble in its matrix is reduced by the dilution of the oscillators (see Fig. 20.3).

In Fig. 23.27 we show data from various authors for the exchange splitting of the excitons in CdSe quantum dots as a function of the dot radius (see also [98G2, 98G3]). Similar data are found for Cu and Ag halogenid dots [00G1, 03G1].

The size dependence of QD electron–hole pair states is demonstrated in Fig. 23.28a, b. The data were taken on colloidal CdSe dots in an organic solvent. The energy of the lowest electron–hole pair state as a function of the dot radius, and the energy spacing to higher states as a function of the energy of the lowest state are in good agreement to theoretical expectations.

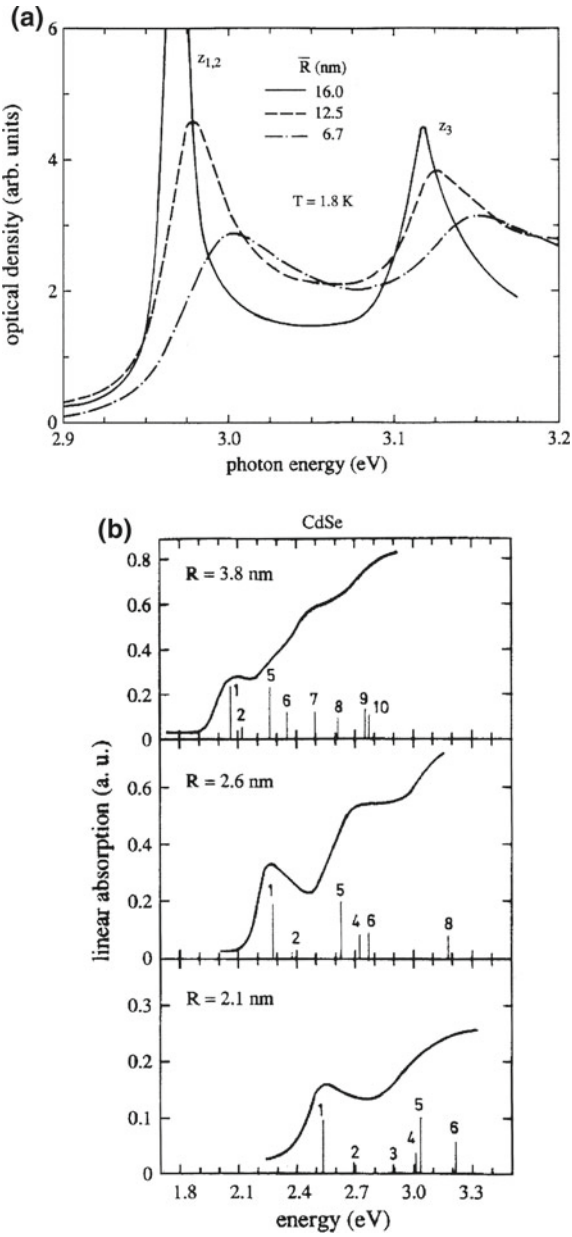


Fig. 23.25 A set of absorption spectra of CuBr quantum dots with different average radii $\bar{R}$ (**a**) (reprinted with permission from [88H1] ©1988 The Optical Society of America) and a set of measured absorption spectra together with the calculated positions of the optical transitions in CdSe QDs (**b**) (reprinted with permission from [93E1] ©1993 Springer Nature)

The absolute photo-luminescence quantum efficiency of colloidal QDs can be rather high, reaching values of close to 100% [13C1]. *Single-dot spectroscopy* reveals that the emission line shows *spectral diffusion* and *fluorescence intermittency* or *blinking* [96N2, 00N1, 01S1, 06G2, 13C1]. The former effect is attributed to the ionization of carriers from the dot and their capture at the interface or in the surroundings of the dot. Also fluctuations of the dielectric environment or polarization due

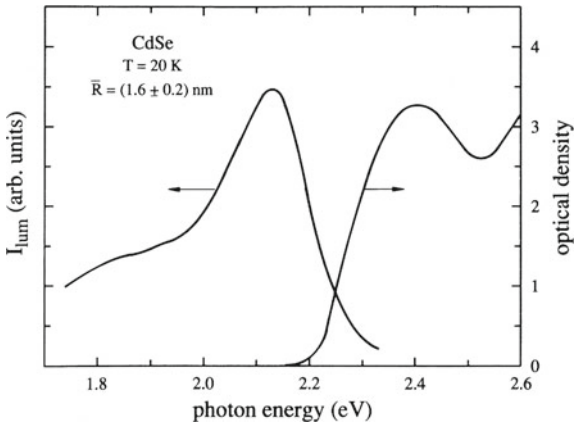


Fig. 23.26 The Stokes-shift between the absorption and emission spectra of CdSe quantum dots (reprinted with permission from [93W3] ©2004 John Wiley and Sons)

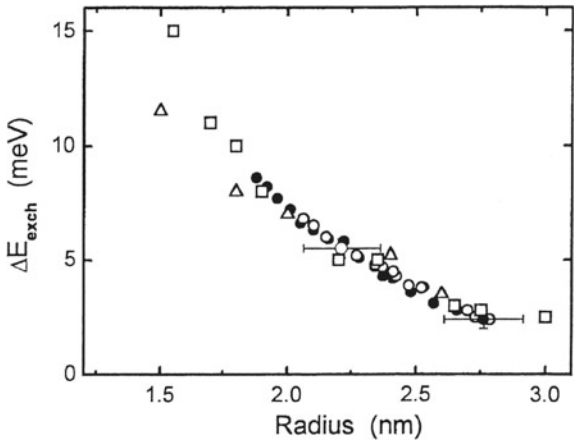


Fig. 23.27 Exchange splitting as a function of the exciton ground state in CdSe dots as a function of the dot radius (Data from various authors, reprinted with permission from [96W1] ©1996 American Physical Society https://doi.org/10.1103/PhysRevB.54.1506)

to surrounding ligands can occur. These result is a quantum-confined Stark effect, caused by the respective electric fields. Carrier transfer to the surface can bring the dots into a non-luminescent state. Removal of the carrier by optical or thermal activation brings the QD back to the original state.

InGaAs-based self-assembled quantum dots are typically considered to have lens-shaped (in-plane rotational) symmetry (D_{2d}). The ideal shape of truncated pyramids can be achieved for large dots, but even these dot never consist of pure InAs due to interdiffusion and segregation effects [07P1]. Since they have interesting applications in spin storage or as single-photon emitters one often focuses on the fundamental optical transitions in the strong confinement regime. Here one uses a simplified notation for exciton states (see, e.g., [02B1]). It starts in first approximation (neglecting heavy-hole–light-hole mixing) with the single-particle basis of the heavy hole with $J_h = 3/2$, $J_{h,z} = \pm 3/2$ and the electron with $S_e = 1/2$, $S_{e,z} = \pm 1/2$. The electrons and holes form two dark ($M = \pm 2$) and two bright ($M = \pm 1$) exciton states with $M = S_{e,z} + J_{h,z}$ being the angular momentum projections. The exciton basis is then termed $|-1\rangle, |+1\rangle, |-2\rangle, |+2\rangle$ (Fig. 23.29).

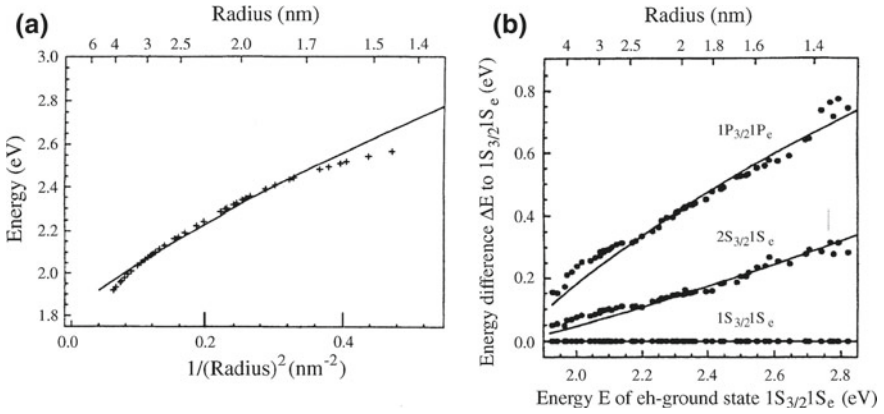


Fig. 23.28 Measured and calculated energies of the lowest electron–hole pair state in CdSe as a function of the dot radius (**a**) and the splitting of higher states as a function of this energy (**b**) (reprinted with permission from [96N1] ©1996 American Physical Society http://doi.org/10.1103/PhysRevB.53.16338)

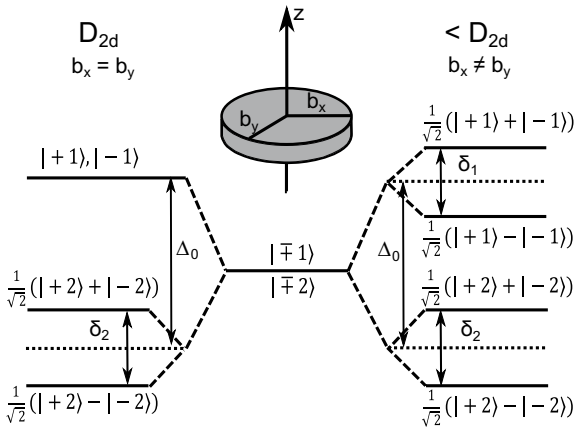


Fig. 23.29 Excitonic fine structure in lens-shaped (left side) and slightly elongated (right side) QDs; the notation of the states and splittings follows [02B1] and is explained in detail in the main text

Due to the exchange interaction and further reduction of dot symmetry the exciton states acquire a distinct fine-structure which determines the optical spectrum of single dots (see Fig. 23.29) [02B1]. In case of lens-shaped dots (symmetry D_{2d}) short-range and long-range exchange together result in a splitting (Δ_0) of dark and bright states. The off-diagonal element δ_2 of the electron–hole exchange Hamiltonian splits the dark doublet while the bright states are still eigenstates of the exchange Hamiltonian and thus degenerate.

In case of dots which are elongated in one planar direction (i.e., with symmetry $< D_{2d}$) the angular momentum is not a good quantum number any longer which now also mixes the $|-1\rangle$ and $|+1\rangle$ states. The long-range exchange interaction (δ_1) splits the hybridized bright states in transverse and longitudinal components while the dark states stay unaffected. The experimentally observed fine-structure splitting of the bright doublet actually depends on the size of the quantum dot and can even be inverted. The later finding results from the interplay of the exchange interaction and

piezoelectric effects due to asymmetric in-plane strain. The elongation of the dot distorts the electron and hole wavefunctions in the same crystallographic direction while the piezoelectric field distorts the wavefunctions in orthogonal directions [05S1].

Experimental Techniques: μ-Photoluminescence and SNOM

A typical density of InGaAs quantum dots could be 10^{11} dots/cm^2. How can one address individual dots in such a large ensemble?

A possible approach is to deposit metal shadow masks with apertures of the size of 100–200 nm on top of the sample. This reduces the number of dots which are addressed simultaneously to some 10. The quantum dots are then excited by illumination with a focused light source (photo-luminescence) or by an electron beam (cathodo-luminescence).

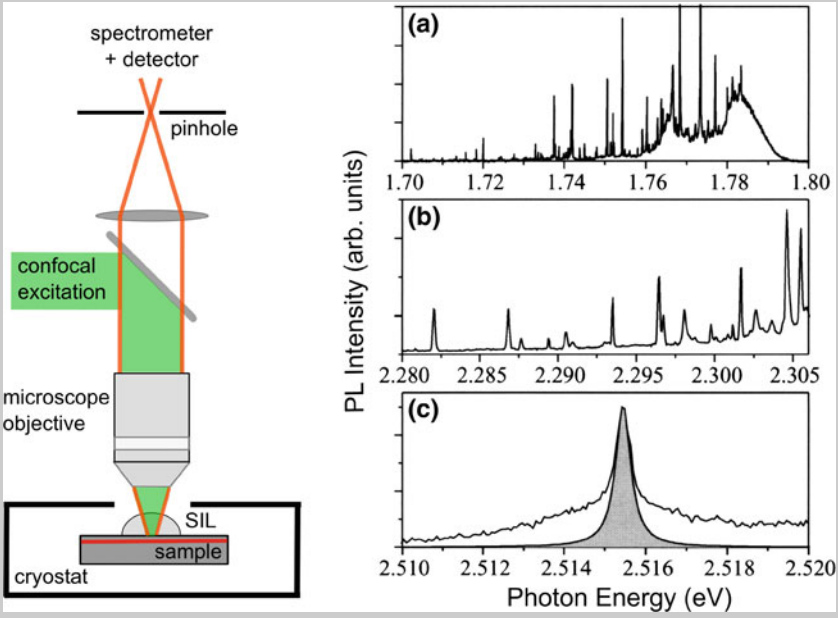


Fig. nPL.1 Left: SIL-enhanced μ-PL setup. Right: A spectrum of a GaAs/AlAs superlattice measured at a sample temperature of 22 K (**a**); Spectrum of a ZnCdSe/ZnSe quantum-dot sample (**b**); Details of an individual emission line of the ZnCdSe/ZnSe quantum-dot sample measured at a sample temperature of 60 K. The shaded area shows a Lorentzian fit to the central part of the peak (**c**) (reprinted with permission from [03M2] ©2003 AIP Publishing https://doi.org/10.1063/1.1567035)

In order to isolate small numbers of QDs one can also fabricate separated mesas or pillars with diameters of some 100 nm from planar samples. For this one has to define protective masks by lithography followed by an etching process leaving the desired structures in the semiconductor sample.

A useful strategy to perform spectroscopy on a small subset of QDs is to use microscopic imaging techniques to excite and/or read-out the dots. This method is called *micro-photoluminescence* (μ-PL) [05K1]. Best spatial resolution is achieved in a confocal arrangement where excitation and detection spot are defined by the same microscope objective and including a pinhole before the detector for spatial filtering [96W5]. The positioning of a solid immersion lens (SIL) right on the sample surface further enhances the resolution to a fraction of the light wavelength (about 0.4λ) and strongly enhances the collection efficiency of the luminescence due to reduction of the total internal reflection at the sample surface [03M2].

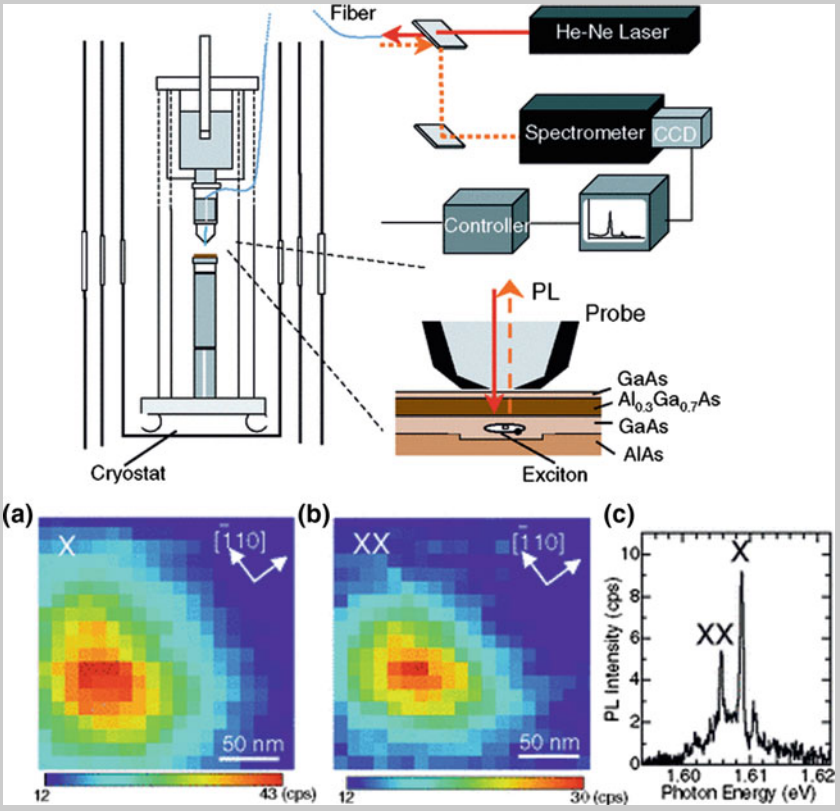


Fig. nPL.2 Upper panel: Schematic of a low temperature scanning near-field optical microscope (SNOM) and setup of a near-field photoluminescence (PL) measurement on a narrow GaAs quantum well incorporating interface quantum dots (IQDs). **Lower panel**: High-resolution PL intensity maps of the exciton (**a**) and the biexciton (**b**) recorded for the respective emission line in the spectrum shown in (**c**) (adapted with permission from [03M3] ©2003 American Physical Society http://doi.org/10.1103/PhysRevLett.91.177401)

All the above methods can address individual QDs or islands due to the fact that these entities have a significant size, shape and composition distribution. As a consequence it possible to isolate emission lines from individual dots in the emission. The spectral resolution of such setups is typically by far not sufficient to determine the homogeneous linewidth of QDs at low temperatures (few μeV, see above). But it is sufficient to separate various excitonic lines like neutral and charged excitons and/or biexcitons. Line-shape fits allow the determination of a line maximum or center of gravity (and thus shifts of such lines upon change of experimental parameters) with a precision of several μeV [00L2].

A *scanning near-field optical microscope (SNOM)* uses a fiber probe for local excitation or light collection. The tip of a tapered single-mode fiber is covered by Au leaving a central aperture of 100 μm or less. The spatial resolution is given by the aperture and the distance between fiber probe and emitting quantum dot. Spatial resolution of 30 nm has been achieved for excitonic localization centers in GaAs quantum wells (interface quantum dots IQDs) covered by a thin barrier/capping layer [03M3].

Theoretical interpretation of the experimentally determined wave-function maps requires consideration of the specific excitation/collection mode of the SNOM and the spatial resolution. For the mapping of biexciton wave-function further corrections have to be applied [05R1]. Near-field condition actually allows to detect excitonic states which are dipole-forbidden in the far-field [04H1]. For the mapping of the excitonic wave function in quantum dots by optical near field microscopy see also, e.g., [01C2, 01G1].

It is interesting to investigate the polarization of the photo-luminescence (PL) from the discussed types of QDs in the strong-confinement limit. Such experiments (see Fig. 23.30) on single dots show that the fine-structure leads to well-defined polarization states which can be tuned in an applied magnetic field. For dots with full axial symmetry (D_{2d}) parallel to the magnetic field B (dot 1) one finds a unpolarized PL which splits into a circularly polarized doublet due to the Zeeman effect (see Sect. 24.1). In slightly elongated dots (symmetry $< D_{2d}$, dot 2) one finds a linearly polarized doublet due to the exchange splitting at zero field. For increasing B field the Zeeman effect restores the axial symmetry which results again in a circularly polarized doublet. The later can split up into a quadruplet for dots where the symmetry is even more reduced (dot 3) which allows for an appearance of dark states in the emission [02B1].

The negative trion in strong confinement has no exchange splitting of bright or the dark doublet so that the $|\pm 1\rangle$ and $|\pm 2\rangle$ are degenerate each [02B1, 02F1]. Due to the reduced relaxation between spin states in this case it is interesting for application of QDs as spin-storage elements.

The linewidth of excitonic emission from single dots is often broadened by spectral diffusion. Still it has been found by spectroscopy with high spatial resolution on both

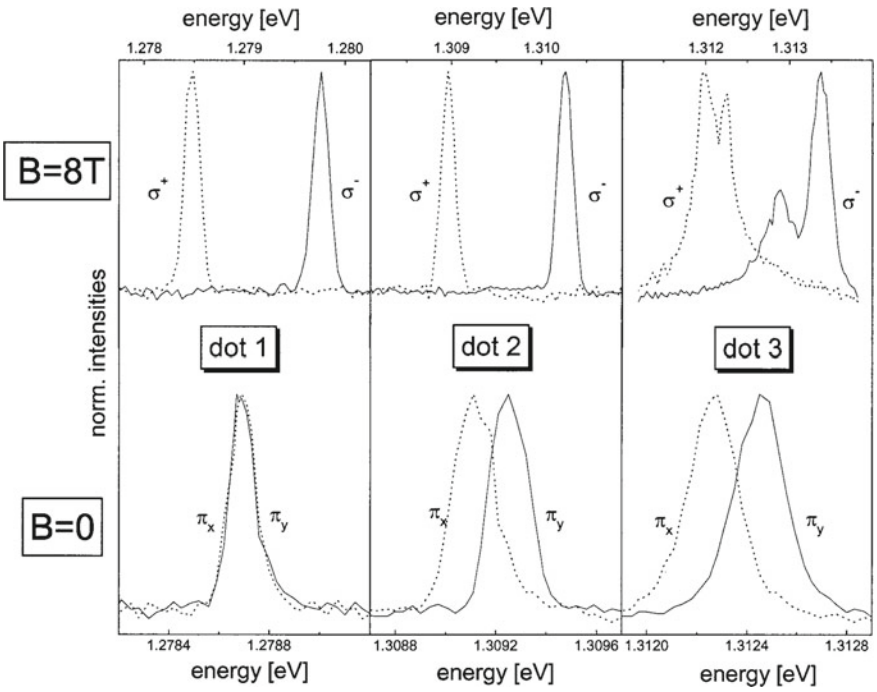


Fig. 23.30 Polarized photoluminescence spectra of three different $In_{0.60}Ga_{0.40}As/GaAs$ self-assembled single quantum dots (the several panels), each representing a class of dots with a specific fine structure pattern at B = 0 (lower traces) and at 8 T (upper traces). The zero-field emission has been analyzed with respect to its linear polarization, and for the high-field emission the circular polarization has been detected (reprinted with permission from [02B1] ©2002 American Physical Society http://doi.org/10.1103/PhysRevB.65.195315)

II–VI and III–V quantum dots, that at low temperatures the emission line consists of a zero-phonon line which in absence of spectral diffusion is homogeneously broadened by the exciton radiative decay (pure dephasing). This line sits on a broad background which results from the interaction of the exciton with acoustic phonons [01B1, 06R1]. These results are consistent with the determination of the dephasing time from four-wave mixing experiments on single InGaAs dots [01B2]. The homogeneous linewidth for temperatures of few Kelvin is found to be few μeV corresponding to several hundred picoseconds excitonic life time. Narrow excitonic emission lines are also observed from single excitons localized in well-width fluctuations in GaAs/AlGaAs quantum wells (interfacial quantum dots (IQDs)) [96G1, 03M3].

There have been significant efforts to achieve efficient luminescence from Si due to the obvious advantages of integration of respective opto-electronic elements into Si technology. The idea was to relax the momentum conservation for the indirect optical transitions due to confinement of the electronic states in space. But, efficient optical emission was not realized till the the introduction of *porous Si* [84P1, 90C3,

97C1]. Porous silicon is prepared by *anodic oxidation* of doped Si, e.g., in HF-ethanol electrolytes [97C2]. The result is a porous layer on the Si chip, which contains nanocrystals and nanorods with dimensions of a few to several nm. These samples show luminescence in the visible spectrum from the red-orange to the blue region depending on the preparation and excitation conditions.

The origin of the bright luminescence has been controversially discussed in a vast amount of papers. Three main models have been proposed. One idea to explain the blue shift and the enhanced intensity of the emission compared to bulk Si is by the quantum confinement effect [97C2]. The second model attributes the origin of the emission to siloxene $(Si_6O_3H_6)_n$ or to Si-oxihydrides or sub-oxides and other Si-related compounds, which may form during the etching process in the porous layer, partly even in the form of surface states or traps at the nanocrystals [95D1].

The third model combines the effect of confinement, which influences the band gap, with the new Si-related compounds as well as surface passivation by Si-H bonds. The energy position of the emission follows the shift of the gap energy while the luminescence yield is enhanced by the passivation [93K1, 93D1, 99W1, 16Y1].

We want to close this section by mentioning another interesting phenomenon. This is the fact that a single quantum dot can emit only one photon at a time in the recombination of a single electron–hole pair. It then has to be re-excited again before it can emit another photon. As a consequence, the photons emitted from a single dot are anti-correlated. This photon anti-bunching, correlation or entanglement of photons emitted from QDs containing two (or even more) electron–hole pairs or coherent control of QD states will be addressed in detail in the second volume of this textbook.

Problems

23.1. In a quantum well with infinitely high barriers one expects an energy level scheme that depends quadratically on the quantum number n_z. Check if a similar relation holds for the (few) bound exciton states in wells with finite barriers. Inspect, e.g., data for the system GaAs/AlGaAs.

23.2. Make a qualitative sketch of the conduction and valence band edges of an ideal, undoped, type-I quantum well, of a barrier and a well-doped sample, as well as of a type-II MQW with a slight homogeneous n-doping.

23.3. Make a rough sketch of the absorption spectrum for the band-to-band transition with and without Coulomb interaction of a direct gap, dipole-allowed bulk semiconductor and for quantum wells, wires and dots.

23.4. Calculate the first few energy levels and wave functions in a spherical and a cubic dot with infinitely high barriers.

23.5. Try to make a simple estimate of the radius of an ideal quantum dot at which the quantization energy and the Coulomb energy are equal.

References

[66S1] M. Shinada, S. Sugano, J. Phys. Soc. Jpn. **21**, 1936 (1966)
[80P1] J.J. Le Pore, J. Appl. Phys. **51**, 6441 (1980)
[81M1] R.C. Miller et al., Phys. Rev. B **24**, 1134 (1981)
[83G1] A.C. Gossard, *Thin Films, Preparation and Properties* (Academic, New York, 1983)
[83K1] D.A. Kleinman, Phys. Rev. B **28**, 871 (1983)
[84G1] R.L. Greene, K.K. Bajaj, Phys. Rev. B **29**, 1807 (1984)
[84P1] C. Pickering et al., J. Phys. C **17**, 6535 (1984)
[84S1] L. Schultheis, K. Ploog, J. Lumin. **31–32**, 394 (1984)
[84S2] L. Schultheis, K. Ploog, Phys. Rev. B **30**, 1090 (1984)
[86A1] Z.I. Alferov et al., JETP Lett. **43**, 569 (1986)
[86D1] G.H. Doehler, IEEE J. QE-**22**, 1682 (1986)
[86D2] G.H. Doehler et al., Solid State Commun. **57**, 563 (1986)
[86S1] H.M. Schmidt, H. Weller, Chem. Phys. Lett. **129**, 615 (1986)
[88C1] S. Charbonneau, M.L.W. Thewalt, E.S. Koteles, B. Elman, Phys. Rev. B **38**, 6287 (1988)
[88H1] F. Henneberger et al., Appl. Phys. B **46**, 19 (1988)
[88K1] M. Kohl et al., Phys. Rev. B **37**, 10927 (1998)
[88K2] E.S. Koteles, J.Y. Chi, Phys. Rev. B **37**, 6332 (1988)
[88K3] P.S. Kop'ev et al., J. Lumin. **40–41**, 747 (1988)
[88K4] Y. Kayanuma, Phys. Rev. B **38**, 9797 (1988)
[88U1] I.N. Uraltsev et al., Phys. Status Solidi (b) **150**, 673 (1988)
[89K1] M. Kohl et al., Superlattices Microstruct. **5**, 235 (1989)
[89K2] E. Kapon, D.M. Hwang, R. Bhat, Phys. Rev. Lett. **63**, 430 (1989)
[89K3] K. Kempa et al., Phys. Rev. B **40**, 8385 (1989)
[89O1] M. Olszabier et al., Phys. Rev. Lett. **62**, 2997 (1989). Surf. Sci. **228**, 123 (1990). J. Lumin. **45**, 186 (1990)
[89S1] S. Schmitt-Rink, D.S. Chemla, D.A.B. Miller, Adv. Phys. **38**, 89 (1989)
[90A1] L.C. Andreani, F. Bassani, Phys. Rev. B **41**, 7536 (1990)
[90A2] A. Alexandrou et al., Phys. Rev. B **42**, 9225 (1990)
[90C1] R. Cingolani et al., Phys. Rev. B **42**, 2893, 3909 (1990)
[90C2] R. Cingolani, L. Tapfer, K. Ploog, Appl. Phys. Lett. **56**, 1233 (1990)
[90C3] L.T. Canham, Appl. Phys. Lett. **57**, 1046 (1990)
[90G1] E.O. Göbel, K. Ploog, Prog. Quantum Electron. **14**, 289 (1990)
[90I1] E.L. Ivchenko et al., Phys. Status Solidi (b) **161**, 219 (1990). Phys. Solid State **36**, 1156 (1994)
[90K1] S.W. Koch, Phys. Bl. **46**, 167 (1990)
[90O1] K. Ogawa, T. Katsuyama, H. Nakamura, Phys. Rev. Lett. **64**, 796 (1990)
[90S1] F.F. So, S.R. Forrest, Y.Q. Shi, W.H. Steier, Appl. Phys. Lett. **56**, 674 (1990)
[90T1] F. Tassone, F. Bassani, L. Andreani, Il Nuovo Cimento **12D**, 1673 (1990)
[91C1] J.H. Collet et al., Phys. Rev. B **43**, 6843 (1991)
[91C2] R. Cingolani, K. Ploog, Adv. Phys. **40**, 535 (1991) and R. Cingolani et al., Phys. Rev. B **42**, 2893, 3909 (1990)
[91E1] A.I. Ekimov, Al.L. Efros, Acta Phys. Polnica A **79**, 5 (1991). A.L. Ekimov, Phys. Scr. **39**, 217 (1991)
[91E2] L. Esaki, in *Highlights in Condensed Matter Physics and Future Prospects*, ed. by L. Esaki (Plenum Press, New York, 1991), p. 55
[91H1] X.-F. He, Phys. Rev. B **43**, 2063 (1991)
[91O1] T. Ogawa, T. Takagahara, Phys. Rev. B **44**, 8138 (1991)
[92K1] H. Kalt et al., Phys. Rev. B **45**, 4253 (1992)
[92K2] S.W. Koch et al., J. Cryst. Growth **117**, 592 (1992)
[92M1] H. Mathieu, P. Lefebvre, P. Christol, J. Appl. Phys. **72**, 300 (1992); Phys. Rev. B **46**, 4092 (1992)

[92M1] S. Malzer et al., Phys. Status Solidi (b) **173**, 459 (1992)
[92O1] D. Oberhauser, Dissertation, Universität Kaiserslautern (1992)
[92O2] K.P. O'Donnel et al., J. Cryst. Growth **117**, 497 (1992)
[92T1] F. Tassone, F. Bassani, Il Nuovo Cimento **14D**, 1241 (1992)
[92T2] F. Tassone, F. Bassani, L. Andreani, Phys. Rev. B **45**, 6023 (1992)
[93C1] P. Christol, P. Lefebvre, H. Mathieu, J. Appl. Phys. **74**, 5626 (1993)
[93C2] D.S. Citrin, Phys. Rev. B **48**, 2535 (1993)
[93D1] C. Delerue, G. Allan, M. Lannoo, Phys. Rev. B **48**, 11024 (1993)
[93E1] A.I. Ekimov et al., J. Opt. Soc. Am. B **10**, 100 (1993)
[93G1] S.V. Gaponenko et al., J. Opt. Soc. Am. B **10**, 1947 (1993)
[93H1] H. Haug, S.W. Koch, *Quantum Theory of the Optical and Electronic Properties of Semiconductors*, 2nd edn. (World Scientific, Singapore, 1993)
[93J1] S. Jorda, U. Rössler, D. Broido, Phys. Rev. B **48**, 1669 (1993)
[93K1] F. Koch, M.R.S. Symp, Proc. **298**, 319 (1993)
[93K2] T. Katsuyama et al., Semicond. Sci. Tech. **8**, 1226 (1993)
[93L1] N. Linder et al., Appl. Phys. Lett. **62**, 1916 (1993)
[93W1] W. Wegscheider, Phys. Rev. Lett. **71**, 4071 (1993)
[93W2] U. Woggon et al., Phys. Rev. B **47**, 3684 (1993)
[93W3] O. Wind et al., Adv. Mater. Opt. Electron. **3**, 89 (1993)
[93Y1] F. Yang et al., Phys. Rev. Lett. **70**, 323 (1993)
[94C1] D.S. Citrin, Solid State Commun. **89**, 139 (1994). Phys. Rev. B **51**, 14361 (1995)
[94F1] S. Fafard, Phys. Rev. B **50**, 1961 (1994)
[94G1] M. Gurioli et al., Solid State Commun. **91**, 931 (1994)
[94G2] M. Grün et al., Superlattices Microstruct. **15**, 463 (1994)
[94J1] C. Janowitz et al., Phys. Rev. B **50**, 2181 (1994)
[94L1] C. Luo et al., Superlattices Microstruct. **15**, 221 (1994)
[94L2] W. Langbein et al., Appl. Phys. Lett. **65**, 2465 (1994)
[94L3] W. Langbein al., J. Cryst. Growth **138**, 191 (1994)
[94R1] D.C. Reynolds et al., Phys. Rev. B **50**, 7461 (1994)
[94S1] M.S. Skolnik et al., Phys. Rev. Lett. **73**, 774 (1994)
[95A1] L.C. Andreani, Phys. Status Solidi (b) **188**, 29 (1995)
[95B1] F. Bassani, G. Czajkowski, A. Trediucci, Z. Phys. B **98**, 41 (1995)
[95B2] L. Banyai et al., Phys. Rev. Lett. **75**, 2188 (1995)
[95C1] D.S. Citrin, Phys. Status Solidi (b) **188**, 43 (1995)
[95D1] D. Dimova-Malinovska et al., Thin Solid Films **255**, 191 (1995)
[95G1] A. Gustafsson et al., Appl. Phys. Lett. **67**, 3673 (1995)
[95H1] P.O. Holtz et al., Phys. Rev. Lett. **74**, 4355 (1995)
[95L1] H. Lee et al., J. Appl. Phys. **78**, 6327 (1995)
[95L2] W. Langbein, M. Hetterich, C. Klingshirn, Phys. Rev. B **51**, 9922 (1995)
[95L3] W. Langbein, Dissertation, Universität Karlsruhe (TH) (1995)
[95S1] Y.P. Song et al., Semicond. Sci. Technol. **10**, 1404 (1995)
[95W1] U. Woggon, S.V. Gaponenko, Phys. Status Solidi (b) **189**, 285 (1995)
[96C1] J. Cerne, Phys. Rev. Lett. **77**, 1131 (1996)
[96G1] D. Gammon et al., Science **273**, 87–90 (1996)
[96L1] W. Langbein et al., Phys. Rev. B **54**, 10784 (1996)
[96L3] W. Langbein, H. Gislason, J.M. Hvam, Phys. Rev. B **54**, 14595 (1996)
[96N1] D.J. Norris, M.G. Bawendi, Phys. Rev. B **53**, 16338 (1996)
[96N2] M. Nirmal et al., Nature **383**, 802 (1996)
[96S1] V. Sirinivas, Y.J. Chen, C.E.C. Wood, J. Opt. Soc. Am. **13**, 989 (1996)
[96S2] T. Stroucken et al., Phys. Rev. B **53**, 2026 (1996)
[96S3] M.S. Salib et al., Phys. Rev. Lett. **77**, 1135 (1996)
[96S4] J. Singh et al., Phys. Rev. B **53**, 15909 (1996)
[96W1] U. Woggon et al., Phys. Rev. B **54**, 1506 (1996)

[96W2] U. Woggon, *Dynamics of Optical Excitations in Quantum Dots of Wide-Gap Semiconductors*, in Festkörperprobleme. Advances in Solid State Physics, vol. 23 (1996), p. 175
[96W3] M.U. Wehner et al., J. Opt. Soc. Am. B **13**, 977 (1996)
[96W4] M.U. Wehner, D. Steinbach, M. Wegener, Phys. Rev. B **54**, R5211 (1996)
[96W5] D. Weber et al., J. Opt. Soc. Am. B **13**, 1241 (1996)
[96W5] R.H. Webb, Rep. Prog. Phys. **59**, 427 (1996)
[97C1] R.T. Collins, P.M. Fauchet, M.A. Tischler, Phys. Today **50**(1) 24 and (8) 83 (1997)
[97C2] A.G. Cullis, L.T. Canham, P.D.J. Calcott, J. Appl. Phys. **82**, 909 (1997)
[97M1] Y. Masumoto, K. Sonobe, Phys. Rev. B **56**, 9734 (1997)
[97P1] M.F. Pereira et al., Phys. Status Solidi (a) **164**, 199 (1997)
[97V1] F. Vouilloz et al., Phys. Rev. Lett. **78**, 1580 (1997)
[97W1] U. Woggon, *Optical Properties of Semiconductor Quantum Dots*. Springer Tracts in Modern Physics, vol. 136 (Springer, Berlin, 1997)
[98A1] H. Akiyama, J. Phys. Condens. Matter **10**, 3095 (1998)
[98B1] M. Bayer, S.N. Walck, T.L. Reinecke, A. Forchel, Phys. Rev. B **57**, 6584 (1998)
[98G1] S.V. Gaponenko, *Optical Properties of Semiconductor Nanocrystals* (Cambride Uinversity Press, Cambridge, 2005)
[98G2] P. Guyot-Sinnoest, M.A. Hines, Appl. Phys. Lett. **72**, 680 (1998)
[98G3] F. Gindele et al., Solid State Commun. **106**, 653 (1998)
[98H1] S. Haas et al., Phys. Rev. B **57**, 14860 (1998)
[98J1] L. Jacah, P. Hawrylak, A. Wójs, *Quantum Dots* (Springer, Berlin, 1998)
[98K1] H. Kalt et al., J. Cryst. Growth **184/185**, 795 (1998)
[98K2] M. Kira, F. Jahnke, S.W. Koch, Phys. Rev. Lett. **81**, 3263 (1998)
[98K3] C. Klingshirn, NATO ASI Ser. B **372**, 143 (1998)
[98P1] W. Petri et al., J. Cryst. Growth **184/185**, 320 (1998)
[98Z1] S. Zimmermann et al., Appl. Phys. Lett. **73**, 154 (1998)
[99D1] A. Dinger et al., Semicond. Sci. Technol. **14**, 595 (1999)
[99D2] R. Denschlag, R. von Baltz, Phys. Status Solidi b **215**, 287 (1999)
[99W1] M.V. Wolkin, J. Jorne, P.M. Fauchet, G. Allan, C. Delerue, Phys. Rev. Lett. **82**, 197 (1999)
[00D1] A. Dinger et al., J. Cryst. Growth **214/215**, 660 (2000)
[00G1] O. Gogolin, Phys. Rev. B **62**, 13053 (2000)
[00L1] K. Leosson et al., Phys. Rev. B **61**, 10322 (2000)
[00L2] D. Lüerßen, R. Bleher, H. Kalt, Phys. Rev. **61**, 15812 (2000)
[00M1] G. Malpuech et al., Phys. Rev. Lett. **85**, 650 (2000)
[00N1] R.G. Neuhauser et al., Phys. Rev. Lett. **85**, 3301 (2000)
[00S1] M. Schmidt et al., Appl. Phys. Lett. **77**, 85 (2000)
[00Y1] D.R. Yakovlev et al., Phys. Rev. B **61**, 2421 (2000)
[01B1] L. Besombes, K. Kheng, L. Marsal, H. Mariette, Phys. Rev. B **63**, 155307 (2001)
[01B2] P. Borri et al., Phys. Rev. Lett. **87**, 157401 (2001)
[01C1] X. Chen et al., Phys. Rev. B **64**, 245304 (2001)
[01C2] A. Crottini et al., Phys. Rev. B **63**, 121313 (2001)
[01G1] J.R. Guest et al., Science **293**, 2224 (2001)
[01H1] M.H. Huang et al., Science **292**, 1897 (2001)
[01H2] J. Hu et al., Science **292**, 2060 (2001)
[01K1] C. Klingshirn (ed.), *Semiconductor Quantum Structures, Optical Properties, Part 1, Landolt-Börnstein - Group III Condensed Matter*, vol. 34C1 (Springer, Berlin, 2001)
[01K2] A.V. Kakovin et al., Solid State Commun. **120**, 259 (2001)
[01K3] E. Kurtz et al., J. Cryst. Growth **227/228**, 630 (2001)
[01L1] L. Börnstein, New Series, Group III (Springer, Heidelberg), **34 C1** and **C2**, C. Klingshirn ed. (2001) and (2004), **C3**, E. Kasper, C. Klingshirn eds. (2007)
[01S1] K.T. Shimizu et al., Phys. Rev. B **63**, 205316 (2001)
[01S2] M.H. Szymanska, P.B. Littlewood, R.J. Needs, Phys. Rev. B **63**, 205317 (2001)
[01S3] M. Schmidt et al., Phys. Status Solidi (b) **229**, 643 (2001)
[01T1] S.A. Tarasenko et al., Semicond. Sci. Technol. **16**, 486 (2001)

[01W1] S. Wachter et al., Phys. Status Solidi (b) **224**, 437 (2001)
[02B1] M. Bayer et al., Phys. Rev. B **65**, 195315 (2002)
[02F1] J.J. Finley et al., Phys. Rev. B **66**, 153316 (2002)
[02H1] G.R. Hayes, B. Devaud, Phys. Status Solidi (a) **190**, 637 (2002)
[02K1] J. Krauß et al., Phys. Rev. Lett. **88**, 036803 (2002)
[02R1] A. Reznitsky et al., Phys. Status Solidi (c) 0(1), 280 (2002)
[03C1] M. Combescot, T. Guillet, Eur. Phys. J. B **34**, 9 (2003)
[03G1] O. Gogolin et al., J. Lumin. **102**–**103**, 414, 451 (2003)
[03K1] A. Kavokin, G. Malpuech, *Cavity Polaritons, Thin Films and Nanostructures*, vol. 32 (Elsevier, Amsterdam, 2003)
[03K2] E. Kurtz et al., in *Spectroscopy of Systems with Spatially Confined Structures* (2001). NATO Science Series II, vol. 90 (Kluwer, Dordrecht, 2002), p. 633 and Phys. Stat. Sol. B, **229**, 519 (2002)
[03K3] G. Kocherscheidt, W. Langbein, V. Savona, Phys. Status Solidi (b) **238**, 486 (2003)
[03M1] G. Mannarini et al., Phys. Status Solidi (b) **238**, 494 (2003)
[03M2] S. Moehl, H. Zhao, B. Dal Don, S. Wachter, H. Kalt, J. Appl. Phys. **93**, 6265 (2003)
[03M3] K. Matsuda et al., Phys. Rev. Lett. **91**, 177401 (2003)
[03N1] U. Neuberth et al., Phys. Status Solidi (b) **238**, 494 (2003)
[03P1] V.V. Popov, T.V. Teperik, N.J.M. Horning, Proc. CAOL'2003, IEEE, Part 1, 163 (2003)
[03Z1] R. Zimmermann, E. Runge, V. Savona, Phys. Status Solidi (b) **238**, 478 (2003)
[04C1] S. Chatterjee et al., Phys. Rev. Lett. **92**, 067402 (2004)
[04H1] U. Hohenester, G. Goldoni, E. Molinari, Appl. Phys. Lett. **84**, 3963 (2004)
[04K1] H. Kalt in C. Klingshirn (ed.), *Semiconductor Quantum Structures, Optical Properties, Part 2* , Landolt-Börnstein - Group III Condensed Matter, Vol. 34C2 (Springer, Berlin, 2004)
[04K2] A. Klochikhin et al., Phys. Rev. B **69**, 085308 (2004)
[04P1] H. Priller et al., Phys. Status Solidi (c) **1**(4), 747 (2004)
[05B1] L. Bányai, S.W. Koch, *Semiconductor Quantum Dots* (World Scientific, Singapore, 2005)
[05B2] S. Banerjee, K.A. Shore, J. Appl. Phys. **97**, 123101 (2005)
[05K1] H. Kalt, Lect. Notes Phys. **658**, 51 (2005)
[05L1] N. Le Thomas et al., Phys. Rev. Lett. **94**, 016803 (2005)
[05O1] T. Otterburg et al., Phys. Rev. B **71**, 033301 (2005)
[05R1] E. Runge, C. Lienau, Phys. Rev. B **71**, 035347 (2005)
[05S1] R. Seguin et al., Phys. Rev. Lett. **95**, 257402 (2005)
[05T1] K. Tanaka et al., Phys. Rev. B **71**, 045312 (2005)
[05V1] Z. Vörös et al., Phys. Rev. Lett. **94**, 226401 (2005)
[06G1] A. Gärtner et al., Appl. Phys. Lett. **89**, 052108 (2006)
[06G2] D.E. Gomez, J. Van Embden, P. Mulvaney, Appl. Phys. Lett. **88**, 154106 (2006)
[06H1] R. Hauschild, H. Kalt, Appl. Phys. Lett. **89**, 123107 (2006)
[06K1] Text mainly taken from H. Kalt in *Advances in Spectroscopy for Lasers and Sensing*, *Proceedings of the NATO Advanced Study Institute on New Developments in Optics and Related Fields, Erice, Italy 2005*, ed. by B. Di Bartolo, O. Forte (Springer, Berlin, 2006)
[06R1] S. Rudin, T.L. Reinecke, M. Bayer, Phys. Rev. B **74**, 161305(R) (2006)
[07D1] B. Deveaud (ed.), *The Physics of Semiconductor Microcavities: From Fundamentals to Nanoscale Devices* (Wiley-VCH, New York, 2007)
[07G1] A. Gärtner et al., Phys. Rev. B **76**, 085304 (2007)
[07P1] T. Passow et al., J. Appl. Phys. **102**, 073511 (2007)
[08F1] J. Fallert et al., Opt. Express **16**, 1126 (2008)
[08S1] L. Sun et al., Phys. Rev. Lett. **100**, 156403 (2008)
[09M1] M. Muntwiller, Y. Qingxin, X.-Y. Zhu, J. Electron Spectrosc. Relat. Phenom. **174**, 116 (2009)
[09P1] K. Pradeesh, J.J. Baumberg, G.V. Prakash, Opt. Express **17**, 22171 (2009)
[10H1] H. Hassanabadi et al., Rev. Mod. Phys. **24**, 2931 (2010)

[10K1] C. Klingshirn, B.K. Meyer, A. Waag, A. Hoffmann, J. Geurts, *Zinc Oxide: From Fundamental Properties Towards Novel Applications*. Springer Series in Materials Science, vol. 120 (Springer, Berlin, 2010)
[10R1] H. Rui et al., Appl. Phys. Lett. **96**, 263303 (2010)
[10Y1] D.R. Yakovlev, W. Ossau, *Magnetic Polarons*. Introduction to the Physics of Diluted Magnetic Semiconductors, ed. by J. Kossut, J.A. Gai (Springer, Berlin, 2010)
[12M1] K.F. Mak et al., Nat. Mater. **12**, 207 (2012)
[13B1] T.C. Berkelbach, M.S. Hybertsen, D.R. Reichman, Phys. Rev. B **88**, 045318 (2013)
[13C1] O. Chen et al., Nat. Mater. **12**, 445 (2013)
[13K1] J. Kang et al., Appl. Phys. Lett. **102**, 012111 (2013)
[13M1] B. Mayer et al., Nat. Commun. **4**, 2931 (2013). https://doi.org/10.1038/ncomms3931
[13T1] P. Tonndorf et al., Opt. Express **21**, 4908 (2013)
[14C1] A. Chernikov et al., Phys. Rev. Lett. **113**, 076802 (2014)
[14H1] K. He et al., Phys. Rev. Lett. **113**, 026803 (2014)
[14L1] Y. Li et al., Phys. Rev. B **90**, 205422 (2014)
[14Y1] Z. Ye et al., Nature **513**, 214 (2014)
[15D1] H. Dery, Y. Song, Phys. Rev. B **92**, 125431 (2015)
[15W1] F. Wu, F. Qu, A.H. MacDonald, Phys. Rev. B **91**, 075310 (2015)
[15Z1] B. Zhu, X. Chen, X. Cui, Sci. Rep. **5**, 9218 (2015)
[16B1] D.N. Basov, M.M. Fogler, F.J. Garcia de Abajo, Science **354**, aag1992 (2016)
[16E1] J.P Echeverry et al., Phys. Rev. B **93**, 121107(R) (2016)
[16M1] K.F. Mak, J. Shan, Nat. Photonics **10**, 216 (2016)
[16N1] K.S. Novoselov et al., Science **353**, aac9439 (2016)
[16Y1] D.T. Yan, N.G. Galkin, Proc. SPIE **101761**, 101761W (2016)
[17B1] L.V. Butov, Superlattices Microstruct. **108**, 2 (2017)
[17C1] F. Cadiz et al., Excitonic linewidth approaching the homogeneous limit in MoS_2-based van der Waals heterostructures. Phys. Rev. X **7**, 021026 (2017). https://doi.org/10.1103/PhysRevX.7.021026
[17C2] Y.-J. Chen et al., Nat. Photon. **11**, 431 (2017)
[17D1] S. Dufferwiel et al., Nat. Photon. **11**, 497 (2017)
[17E1] D.K. Efimkin, A.H. MacDonald, Phys. Rev. B **95**, 035417 (2017)
[17L1] T. Low et al., Nat. Mater. **16**, 182 (2017)
[17S1] Z. Sun et al., Nat. Photon. **11**, 491 (2017)
[17Z1] Y. Zhou et al., Nat. Nanotechnol. **12**, 856 (2017)
[18P1] K.D. Park et al., Nat. Nanotechnol. **13**, 59 (2018)
[18S1] G. Scuri et al., Phys. Rev. Lett. **120**, 037402 (2018)

Chapter 24
Excitons Under the Influence of External or Internal Fields

Abstract Application of external fields is a well suited strategy to reveal the properties of excitons in semiconductors. It further leads to new phenomena which are interesting for device applications. The external field can be applied temporally constant [79E1, 85H1] or as systematic perturbation in modulation techniques [69C1, 73S1, 04G1]. We will restrict our discussion here to the presentation of general features and of some selected topics and examples. We consider the influence of magnetic, electric and strain fields on the optical properties of excitons including their continuum states and we proceed again from bulk materials to structures of lower quasi-dimensionality. Mostly these fields are applied from outside, but there are also cases of internal fields, e.g., strain fields due to lattice misfit in epitaxy or piezo-electric fields resulting from such strain.

The implications of the fields are usually treated in perturbation theory of first or second order. Group theory is the ideal tool to know, which matrix elements related to the perturbation vanish and which have finite values. We give two simple examples, how symmetry arguments work (and group theory is essentially based on symmetries):

- A nondegenerate energy level in a system with inversion symmetry, in which parity is a good quantum number, can shift in first order perturbation theory only under the influence of a perturbation with even parity.
- A degenerate level splits under the influence of a perturbation only if either the perturbation reduces the symmetry of the system, e.g. for the $n_B = 1A\Gamma_6$ exciton in CdS [77V1], or if there is an accidental degeneracy only as in the case of the $n_B = 1$ $A\Gamma_1$ and $A\Gamma_2$ excitons in ZnO (see Fig. 24.2).

A short introduction to group theory and its applications to semiconductor optics are given in Appendix B, where also the terms used above will be explained.

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_24

24.1 Excitons and Carriers in Magnetic Fields

24.1.1 Influence of Magnetic Fields on Excitonic Resonances

We start with the influence of a magnetic field $\boldsymbol{B}$ on excitons. Here we have two natural energy scales which have to be compared. One is the excitonic Rydberg energy Ry^* the other the cyclotron energy $\hbar\omega_\mathrm{c} = \hbar(e\boldsymbol{B}/\mu)$, where μ is the reduced mass of the exciton.

The regime

$$\mathrm{Ry}^* \gg \hbar\omega_\mathrm{c} \quad \Rightarrow \quad \gamma = \hbar\frac{eB}{\mu \mathrm{Ry}^*} \ll 1 \tag{24.1a}$$

characterizes the *weak-field limit*. The Coulomb energy dominates and the magnetic field can be treated as a perturbation.

In the *strong-field limit*

$$\mathrm{Ry}^* \ll \hbar\omega_\mathrm{c} \quad \Rightarrow \quad \gamma \gg 1\,, \tag{24.1b}$$

we have to consider first the Landau levels resulting from the quantization of the free particles in reciprocal space related to their motion in the plane perpendicular to $\boldsymbol{B}$, and then the Coulomb energy.

The intermediate regime with $\gamma \approx 1$, which pertains in many semiconductors in typical dc-fields of superconducting or Bitter-type magnets ($\boldsymbol{B}_\mathrm{max} \approx 40\,\mathrm{T}$) is more complicated to describe quantitatively.

It should be mentioned that a sufficiently strong magnetic field also produces quasi one-dimensional subbands similar to the case of quantum wires, i.e., of quasi one-dimensional systems (Sect. 17.4) according to the following argument. The carriers can move freely only parallel to $\boldsymbol{B}$, i.e., in one dimension. In the plane perpendicular to $\boldsymbol{B}$ the motion is quantized to circles in the classical picture resulting in a quantum mechanical description in a harmonic-oscillator-like term level scheme, the so-called *Landau levels*,

$$E\left(n_\mathrm{l}, \boldsymbol{k}_\parallel\right) - E_0 = \left(n_l + \frac{1}{2}\right)\hbar\omega_\mathrm{c} + \frac{\hbar^2\boldsymbol{k}_\parallel^2}{2m_\mathrm{eff}}\,, \tag{24.1c}$$

where n_l is the quantum number of the Landau level and E_0 the band extremum without magnetic field. The density of states is again as in quantum wires for every subband proportional to $(E - E_{n_\mathrm{l},\boldsymbol{k}_\parallel})^{-1/2}$, i.e., one obtains a similar picture as in Fig. 17.1 for vanishing damping with the main difference that the Landau levels are equidistant in energy. Damping washes out the singularities (see below). The selection rules for inter-Landau level transition are within one fan of Landau-levels and for simple, parabolic bands $\Delta n_\mathrm{l} = \pm 1$ and for interband transitions $n_{\mathrm{l,CB}} - n_{\mathrm{l,VB}} = 0$.

We want to treat here the main effects resulting from an applied magnetic field, namely the *diamagnetic shift*, the *Zeeman splitting*, and the appearance of Landau

levels, in a very elementary way and give some examples. References leading deeper into this field are, e.g., [69C1, 73S1, 75C1, 77V1, 80S1, 82L1, 85H1, 97P1, 01K1, 04G1].

If an exciton is situated in a magnetic field the relative motion of electron and hole is deformed by Lorentz forces. In perturbation theory, this deformation can be described by a weak admixture of other states. If we consider the ground-state $n_B = 1, l = 0$, which has an S-like envelope function without angular momentum, we can describe this deformation as a small admixture of $l = 1$ (or P-) envelope states, e.g., via second-order perturbation theory [79E1, 85H1].

The angular momentum resulting from this admixture is proportional to $\boldsymbol{B}$ and is oriented according to Lenz' rule (i.e., the minus sign in (2.1c)) antiparallel to $\boldsymbol{B}$. Since the energy of a magnetic dipole in a magnetic field also increases linearly with $\boldsymbol{B}$ we obtain in total a quadratic, so-called *diamagnetic shift* ΔE_{dia} to higher energies

$$\Delta E_{\text{dia}} = a\boldsymbol{B}^2 \,. \tag{24.2a}$$

The constant a is a material parameter which is proportional to the square of the Bohr radius of the exciton or more general, to its cross-section normal to $\boldsymbol{B}$. The dependence on a_B or n_B and on the material parameters explains that the diamagnetic shift increases with n_B.

A typical value for the A-excitons in CdS ($a_B \approx 2.8\,\text{nm}$) is $a(n_B = 1) \approx 2 \times 10^{-6}\,\text{eVT}^{-2}$, i.e., for 10 T one finds a diamagnetic blue-shift of about 0.2 meV (see [77V1, 85H1] and Fig. 24.3), which is just at the detection limit. For semiconductors with large excitonic Bohr radii like GaAs, the diamagnetic shift is easily observable [74F1]. For the opposite situation realized e.g., in ZnO [82B1] or the Cu-halides it is very difficult to observe with typical stationary $\boldsymbol{B}$ fields in the 20 T range [73S2, 80S1].

Exciton states that at $B = 0$ already have a non-vanishing magnetic moment, which can be aligned relative to $\boldsymbol{B}$, exhibit in addition the *linear Zeeman splitting*. This magnetic moment can come from the spins of electron and hole. For singlet and triplet excitons with S envelope the difference or the sum of electron and hole g-factors enters, in simpliest approximation,

$$\Delta E_z = \pm\frac{1}{2} \mid g_e \pm g_h \mid \mu_B B \tag{24.2b}$$

due to the relative alignment of electron and hole spin. The g-factors of electrons and holes can deviate significantly from two due to influences of the bandstructure or crystal symmetry. Tables are found e.g. in [82B1, 82L2, 85H1, 01L1, 02L1].

For states with $n_B \geqslant 2$ there is an additional contribution from the magnetic moment of the envelope function for $l \geqslant 1$, depending on the orientational (magnetic) quantum number m in Sects. 20.2 and 20.3.

If the states, which are subject to Zeeman splitting are split for some other reasons already for $\boldsymbol{B} = 0$ by an amount Δ, the Zeeman splitting is suppressed until ΔE_z

becomes according to (24.2b) comparable to Δ. This fact can be easily seen by solving (or diagonalizing) a 2×2 Hamiltonian of the following type

$$\begin{pmatrix} E_0 + aB^2 + \frac{\hbar^2 k_\parallel{}^2}{2m} - E & -\frac{1}{2}\mu_B g B \\ -\frac{1}{2}\mu_B g B & E_0 + \Delta + aB^2 + \frac{\hbar^2 k_\parallel{}^2}{2m} - E \end{pmatrix} \tag{24.3}$$

For a quantitative calculation, perturbation theory for (almost) degenerate states has to be used, e.g., in the eightfold space of $n_B = 1$ excitons which can be constructed from the fourfold degenerate Γ_8 valence band and the twofold degenerate Γ_6 conduction band in T_d symmetry. The terms describing the diamagnetic shift appear together with the kinetic energy terms, the singlet-triplet and the longitudinal-transverse splitting in the main diagonal of the resulting 8×8 matrix, while the Zeeman terms, k-linear terms and others contribute to the off-diagonal elements.

At the end of this small introduction we want to mention an additional effect, the *excitonic Stark effect* [61T1, 74O1, 78R1, 98L1, 02K1]. The basic idea is as follows: Assume, that an exciton moves with constant centre of mass velocity $\mathbf{v}$ (or total momentum $\hbar\boldsymbol{k}$, see (20.1) and (20.4)) through a crystal perpendicular to an applied magnetic field $\boldsymbol{B}$. Then we can either state, that the Lorentz forces acting on electron and hole and resulting from the translational motion of the exciton will result in modifications of the excitonic (envelope-) function and of the eigenenergy. Or we can perform a coordinate transformation from the system S, which is fixed in the crystal lattice to a system S' moving with the centre of mass of the exciton. The (relativistic) transformation of $\boldsymbol{B}$ from S to S' results in an electric field $\boldsymbol{E}$, which is perpendicular to $\boldsymbol{B}$ and $\boldsymbol{k}$. This $\boldsymbol{E}$ field influences the exciton via the Stark effect (see Sect. 24.2) explaining the name of the phenomenon.

A more detailed description of the phenomena evoked by a magnetic field, e.g., in terms of a spin formalism or an invariant expansion is beyond the scope of this book. We want to refer the reader to, e.g., [61T1, 79E1, 82B1, 85H1].

In the following sections we want to describe some properties of excitons and carriers in magnetic fields.

24.1.2 Magnetic Fields Applied to Non-magnetic Bulk Semiconductors

In the following we will give some examples for the diamagnetic shift and the Zeeman splitting observed in the magneto-optic spectra of different bulk semiconductors.

We start with the case of the of the dipole allowed $n_B = 1$ AΓ_5-exciton resonance of CdS (Fig. 24.1). With increasing magnetic field a splitting connected with the Zeeman occurs while the diamagnetic shift is hardly visible (see discussion in the previous subsection).

In Fig. 24.2 we give transmission spectra of ZnO in the spectral region of the $n_B = 1$ A- and B-exciton resonances, for the polarization $\boldsymbol{E} \parallel \boldsymbol{c}$. In this orientation

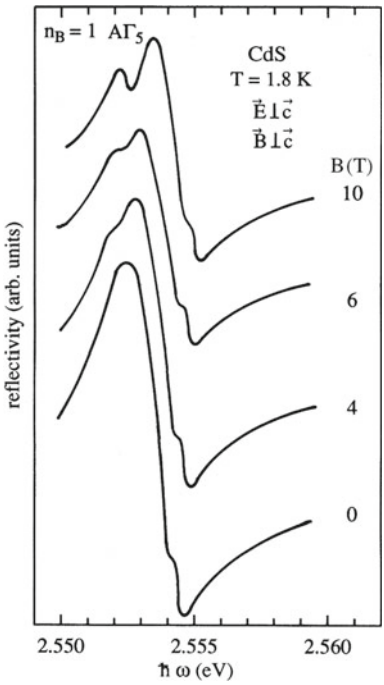


Fig. 24.1 Reflection spectra of the $n_B = 1A\Gamma_5$-exciton resonance for various magnetic fields (reprinted with permission from [82B1] ©1982 American Physical Society http://doi.org/10.1103/PhysRevB.25.7413)

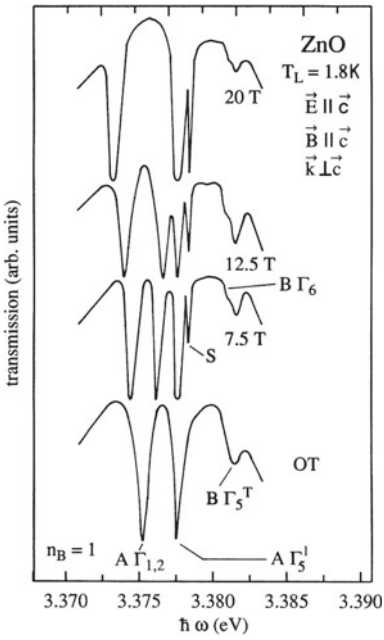


Fig. 24.2 Transmission spectra in the region of the $n_B = 1A$ and B exciton resonances of ZnO for the orientation $\boldsymbol{E} \parallel \boldsymbol{c}$ (reprinted with permission from [82B1] ©1982 American Physical Society http://doi.org/10.1103/PhysRevB.25.7413)

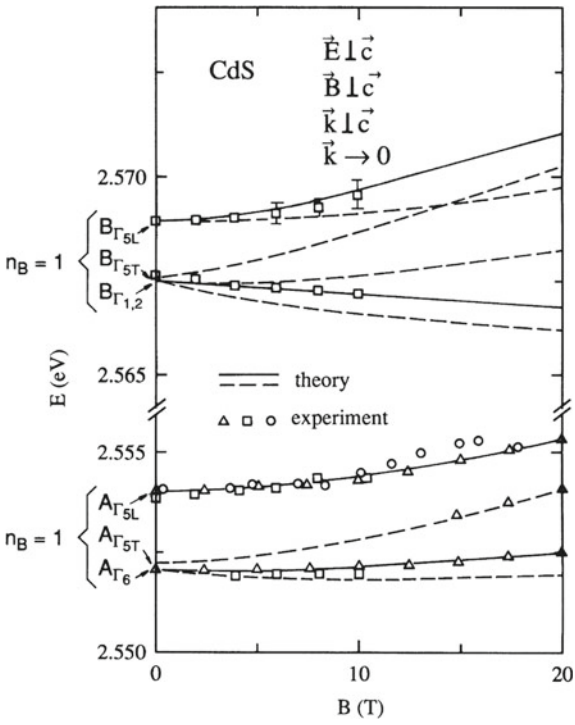


Fig. 24.3 The splitting pattern of the $n_B = 1$A- and B-exciton resonances in CdS with increasing magnetic field (reprinted with permission from [82B1] ©1982 American Physical Society http://doi.org/10.1103/PhysRevB.25.7413)

all transitions have only very weak oscillator strength. Again we observe the Zeeman splitting of the $A\Gamma_{1,2}$ triplet states with increasing $\boldsymbol{B}$, which show at $\boldsymbol{B} = 0$ an accidental degeneracy, and the $A\Gamma_5^L$ state. The narrow spike S is due to an isotropic point, i.e., a photon energy for which the exciton polariton dispersion relations for the polarizations $\boldsymbol{E} \parallel \boldsymbol{c}$ and $\boldsymbol{E} \perp \boldsymbol{c}$ cross or in other words where the ordinary and extraordinary refractive indices are equal. In this situation energy can be transferred from one polarization (here $\boldsymbol{E} \parallel \boldsymbol{c}$) to the other under energy and momentum conservation, leading to the dip in transmission. For a discussion of the g-factors resulting from these data see also [60H1, 02L1, 03C2, 04A1, 04M1, 10K1].

The combined influences of terms linear and quadratic in $\boldsymbol{B}$ is nicely reflected in Fig. 24.3. It depicts the splitting pattern of the $n_B = 1$A- and B-exciton resonances in CdS as a function of the magnetic field for the orientation $\boldsymbol{B} \perp \boldsymbol{c}$, $\boldsymbol{E} \perp \boldsymbol{c}$.

The influence of the magnetic field on the dispersion of exciton polaritons—now called *magneto-polaritons*—is demonstrated in Fig. 24.4. The left panel (a) gives the dispersion relation of the $n_B = 1\Gamma_5$-exciton-polariton resonance in cubic (T_d) ZnTe. The different dispersions for light-hole and heavy-hole excitons is clearly visible. Figure 24.4b shows the situation for finite $\boldsymbol{B}$ in a fairly arbitrary direction. In this situation all degeneracies are lifted and all exciton branches acquire some oscillator strength due to admixtures resulting in a total of eight different exciton and nine polariton branches. This is a rather frightening example for a magneto-polariton

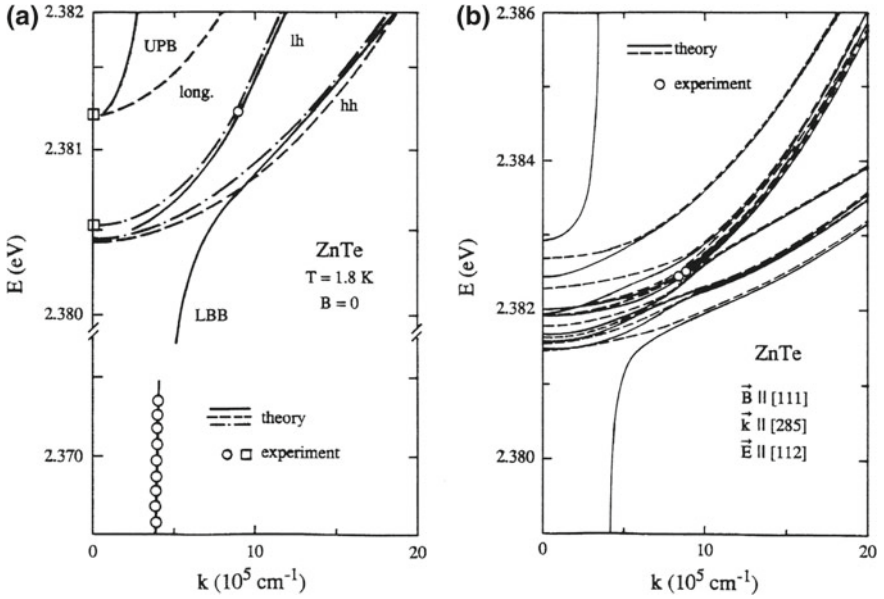


Fig. 24.4 Dispersion relation of the $n_B = 1$ exciton poloriton resonances in ZnTe without (**a**) and with (**b**) an applied magnetic field (reprinted with permission from [85H1] ©1985 Elsevier)

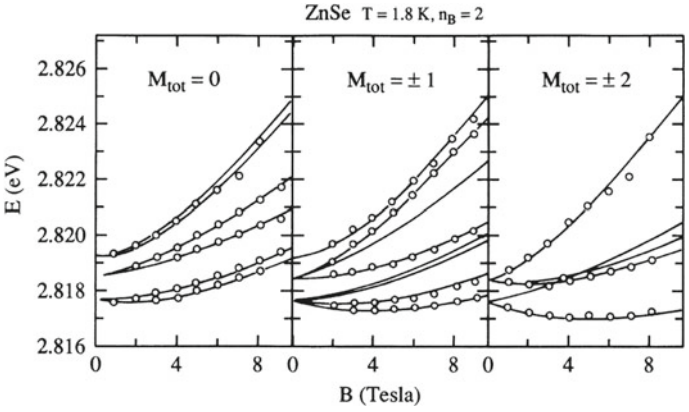


Fig. 24.5 The influence of a magnetic field on the $n_B = 2$ exciton resonances in ZnSe, determined in two-photon absorption spectroscopy (reprinted with permission from [82H1] ©1983 Elsevier)

dispersion teaching us that it is wise to stick to measurement geometries with high symmetry to keep results simple to interpret.

So far we have considered in the examples exclusively the excitonic ground state $n_B = 1$. The magnetic field behavior of the $n_B = 2$, $l = 1$ exciton resonances in ZnSe is now give in Fig. 24.5. There is already a finite splitting for $B = 0$. For higher

fields this splitting increases due to Zeeman terms. In addition, there is a pronounced diamagnetic shift due to the larger radius of the relative motion of electron and hole. The total angular momentum M, consisting of spin and envelope contributions, could be chosen in the two-photon absorption experiments of Fig. 24.5 by using circular and/or linear polarizations of the two overlapping excitation beams.

Further selected examples of magneto-optics comprise the wide gap Cu-halides [73S2, 77V1, 80S1], the insulating alkali halides [86B2] or the normal semiconductor GaAs [74F1]. The dipole forbidden excitons of the yellow series of Cu_2O and the $\boldsymbol{B}$-field induced mixing of ortho and para excitons (see Sect. 21.2.1.2) are treated in [07B1, 08S2]. The magneto-Stark effect has been observed e.g. in CdS [61T1] or in the layered semiconductor GaSe [74O1, 78R1].

Similar splitting patterns as for free excitons can also be observed for *bound exciton complexes* (BEC), where one has to take into account that the $\boldsymbol{B}$-field influences both the initial and the final state, e.g., in a recombination process from A^0X to A^0. Examples of the magnetic field behavior of BEC are given in [79E1, 81B1, 82K1, 86G1, 89G1, 04A1, 04M1, 11C1] or Chap. 8 of [10K1]. Note that it is non trivial to distinguish between a D^0X and a A^0X state from Zeeman splitting alone, because D^0 has an unpaired electron in the ground state and an unpaired hole in the D^0X state, while the situation is just opposite for A^0 and A^0X. The thermalization behavior between the various Zeeman split components in luminescence or absorption may give additional information. See e.g. [81B1] or Chap. 8 of [10K1] and references given therein.

For more recent results on the $D^0 \rightarrow D^0X$ transition in GaAs in a magnetic field resulting in a three level Λ system see [05K1]. Transitions within the hydrogenic series of states in the D^0 of GaAs by THz spectroscopy in a magnetic field are detailed in [05D1].

At the beginning of this section we noted that the observation of *Landau levels* might be difficult, since for most semiconductors the limit $\gamma \gg 1$ can be reached only for fields $\boldsymbol{B} \gtrsim 10^2$ T, which, if at all, are usually available only in pulsed form.

There is, however, a way to overcome this difficulty by investigating the continuum states, where the Coulomb interaction still influences the oscillator strength but where the motion of electron and hole is almost free. To observe the Landau levels, their broadening must be smaller than the cyclotron energy as in micro-wave experiments, in other words $\omega_c T_2 \gtrsim 1$ where T_2 is the phase-relaxation time. Bulk samples with direct, dipole-allowed band-to-band transitions are usually opaque in this region so that reflection spectroscopy is the appropriate tool.

In Fig. 24.6 we give as an example the reflectivity of ZnO at the onset of the continuum states of the *C*-exciton series. The Landau-level structure becomes obvious for the highest *B*-fields. The observation of the Landau levels allows the sum of electron and hole masses to be determined. If the electron mass is known, e.g., from cyclotron absorption by *n*-type materials, the hole mass can then be determined.

In semiconductors with relatively small exciton binding energies like GaAs, the magnetic field may shift states into the continuum, resulting in Fano-interference-type lineshapes. The quantization of the continuum leads to a reduction of this effect. For some examples see, e.g., [95S1, 97B1].

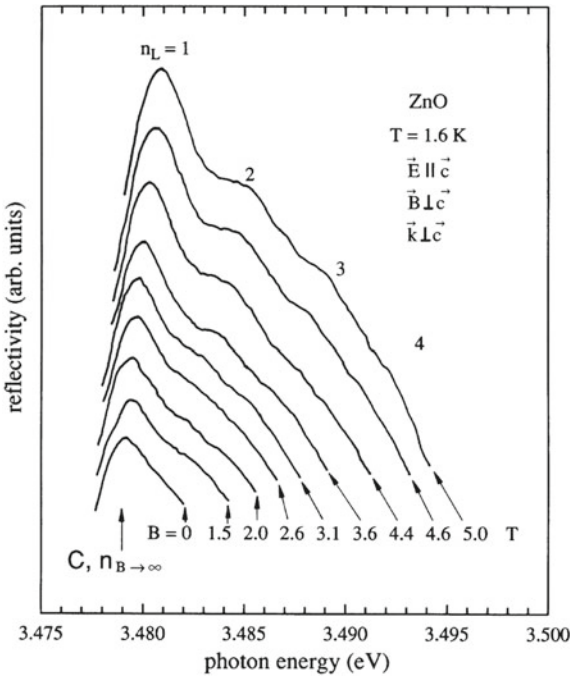


Fig. 24.6 The reflectivity in the spectral region of the exciton continuum in ZnO for various magnetic fields (reprinted with permission from [73H1] ©2006 John Wiley and Sons)

Until now we concentrated on shifts and splittings of eigenenergies as a function of the magnetic field. Via Kramers–Kronig relations (see Appendix A) these effects influence also the refractive index $n(\omega)$, i.e. the polariton dispersion (Fig. 8.1). If it is different for left and right circular polarized light $\sigma^{\pm}$ and if $\boldsymbol{k}$ is oriented parallel to $\boldsymbol{B}$, one can observe the Faraday effect. This is a rotation of the polarization plane of linearly polarized light by an angle α which grows linearly with magnetic field $\boldsymbol{B}$ and the distance l over which the light travels through the medium

$$\alpha = V(\omega) B l \,. \tag{24.4}$$

The proportionality constant V is called Verdet-constant. It depends on ω and increases when ω approaches a resonance. For an example of the observation of the Faraday effect in a semiconductor see e.g. [73S2] for CuCl and for semimagnetic semiconductors Sect. 24.1.3.

24.1.3 Magnetic Fields Applied to Bulk Diluted Magnetic Semiconductors

Diluted magnetic or *semi-magnetic semiconductors* (DMS or SMSC) are semiconductors containing magnetic ions. The coupling of their magnetic moment to the

carriers and excitons of the semiconductor strongly affects the optical properties. Some reviews and data are found in [78G1, 82L2, 88G2, 88S1, 91O1, 92G2, 91S1, 92Y1, 94G1, 95M1, 96Y1, 01D1, 01O1, 02S1, 03D1, 03D2, 04H1] and in Chap. 6 of [03O1] for DMS lead chalcogenides. While the older articles concentrate more on bulk materials, the more recent ones give general results for quantum structures. For the latter see Sect. 24.1.4.

The DMS are generally II–VI or III–V semiconductors doped with magnetic ions up to several tens of percent. This means that they can also be considered as alloys for higher concentrations, with the corresponding consequences for composition dependencies for phonons, the band gap, exciton binding energies and localization, as outlined in Sects. 13.1, 18.2, and 22.2. The doping atoms must carry, by definition, spin and magnetic moment. The most widely used one is Mn^{2+} with a half-filled 3d shell but also the influences of Fe, Co, Ni, V, Cr, Eu or Gd are investigated. These atoms sit ideally on the metal or cation site of the host. The most widely used host materials are narrow and wide gap II–VI materials like $Hg_{1-x}Mn_xTe$, $Cd_{1-x}Mn_xTe$ or $Zn_{1-x}Mn_xSe$. Also ZnO:V and ZnO:Co have been investigated in addition to III–V compounds like GaAs:Mn. Since Mn^{2+} acts on a Ga site as an acceptor, GaAs:Mn samples are metallic without compensating co-doping.

Most of the binary DMS compounds show antiferromagnetic ordering like MnTe and $Cd_{1-x}Mn_xTe$ for concentrations x down to about 0.35. Some others are ferromagnetic like $Ga_{1-x}Mn_xAs$ or $Ga_{1-x}Mn_xN$, possibly Zn_{1-x} Mn_xO or ZnO:V and ZnO:Co [98O1, 00D1, 02D1, 02S2, 02S3]. However, the origin of the observed, very weak signals in DMS II–VI compounds is still discussed controversially and partly attributed to clusters or precipitates and even to the substrate [07D1, 07K1]. See for ZnO based DMS alloys also the discussion in Chap. 10 of [10K1]. The ferromagnetism can be influenced by the concentration of free carriers, e.g, holes in GaAs:Mn or in oxide semiconductors [06P1]. Their density can be decreased by codoping with a donor. Similarly codoping with a donor like Ga or Al increases the electron density in ZnO:V or ZnO:Co. Some theoretical models predict for Zn_{1-x} Mn_xO ferromagnetism only for p-type doping levels which are far beyond experimental reach [00D1]. At low concentration the DMS alloys are paramagnetic with very high saturation magnetization and can exhibit spin glass behavior at low T.

We now start to review more of the specific semi-magnetic properties of bulk DMS. Since close pairs of magnetic ions couple generally antiferromagnetic [04H1], they do not contribute to the strong para—nor to the possible ferromagnetism of DMS. A nominal concentration $x_{nom} = 0.15$ in $Zn_{1-x}Mn_xSe$ may well result in an effective concentration $x_{eff} = 0.05$ deduced, e.g., from saturation magnetization [06H1]. The paramagnetism, the giant Zeeman splitting (see below) or the ferromagnetism go with increasing x_{nom} through a maximum and disappear when antiferromagnetic ordering dominates.

The main point is that the exchange interaction occurs between the s- and/or p-type conduction and valence bands of the host and the 3d orbitals of the magnetic ions, known as sp-d exchange. A free or bound carrier can polarize the magnetic moments of the surroundings, e.g., Mn^{2+} ions. This fact results in a Giant Zeeman splitting and large effective g-factors for both the free or bound carriers and for the excitons. The

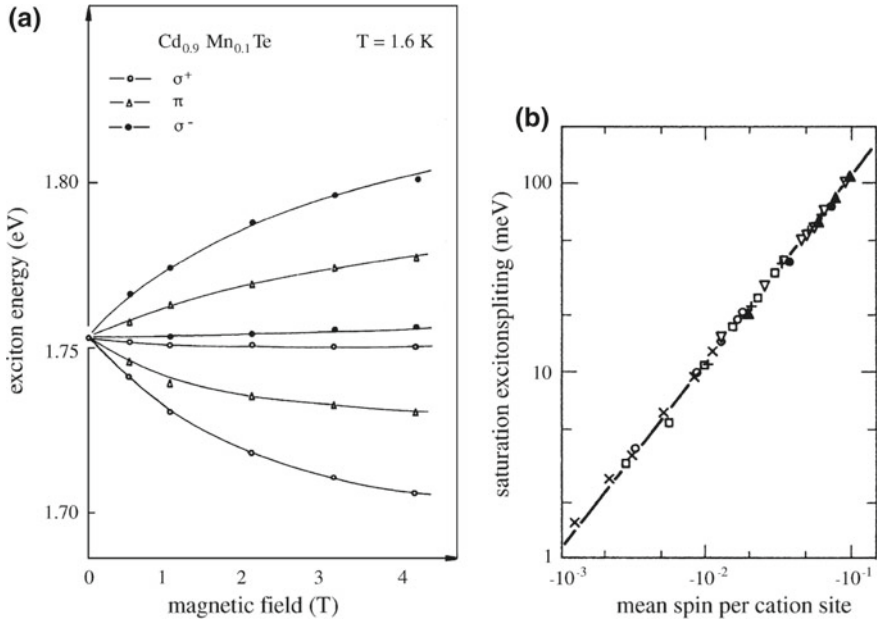


Fig. 24.7 Zeeman splitting of the exciton of $Cd_{0.9}Mn_{0.1}Te$ as a function of magnetic field in the Faraday configuration, i.e., $\boldsymbol{k} \parallel \boldsymbol{B}$ (**a**). Saturation Zeeman splitting per average Mn^{2+} spin per cation site (**b**) (reprinted with permission from [88G1] ©1988 Elsevier)

effective g-factors can reach values up to 100 and the Zeemann splitting saturates at values between 10 and 100 meV for fields of a few T.

Both the splitting values and the g-factors are thus by one or two orders of magnitude larger than in nonmagnetic semiconductors. We give an example in Fig. 24.7 that can be compared, e.g., to Figs. 24.1, 24.2 and 24.3. These large values of the Zeeman splitting are also held responsible for the large values of *Faraday* and *Kerr rotation* observed in transmission and reflection geometries, respectively. For examples see, e.g., [78G1] and for magnetic circular dichroism [08A1].

Free and even more bound carriers can form so-called *magnetic polarons*. This is a carrier surrounded by a magnetization cloud of oriented Mn^{2+} moments in a similar way as a usual carrier can be surrounded by a phonon cloud forming a polaron (see Sect. 14.6). The formation time of a magnetic polaron after creation of a carrier (e.g., by optical excitation) is finite and in the range of 100 ps or longer. The formation of the magnetic polaron cloud lowers the energy of the carrier. This energy decreases with increasing temperature and increasing external magnetic field [92Y1, 96Y1].

In narrow gap DMS one observes with increasing magnetic field well-resolved *Landau-level fans* [91O1]. Mn^{2+} ions in II–VI DMS show an internal transition resulting in a luminescence band around 2.2 eV, i.e., in the yellow spectral range and absorption features starting at or above this energy when the band gap, or to phrase

it better exciton energies, are situated above theses transitions. This is the case for $Cd_{1-x}Mn_xTe$ approximately for $x > 0.4$ [84L1].

Some of the DMS show a ferromagnetic hysteresis loop, partly up to room temperature [00D1, 06K1]. As mentioned already above, in some cases like ZnO:X with X = Mn, Co, V... there is an ongoing discussion whether this ferromagnetism is of intrinsic nature or due to some ferromagnetic clusters or precipitates [05O1, 06K1] or Chap. 10 of [10K1].

The ferromagnetic magnetization M_{DMS} is often given in so-called electromagnetic units (emu) or emu/g. Here the relation holds

$$1\,\text{emu} = 10^{-3}\,\text{A/m}. \tag{24.5}$$

Since typical numbers are in the range of 10^{-3} emu or even below (24.5) results with (2.1f) in magnetic flux densities B_{DMS} of the order of $B_{DMS} = \mu_0 M_{DMS} \approx 10^{-12}\,\text{Vs/m}^2 = 10^{-12}\,\text{T}$.

Frequently the saturation magnetisation is, apart from its sign, comparable to the diamagnetic magnetisation at the same field strength [06P1].

This value is several orders of magnitude below the magnetic field of the earth $B_{earth} \approx 20\,\mu\text{T}$ and even more so below the field of a good commercial ferromagnet $B_{com} \approx 1\,\text{T}$.

24.1.4 Magnetic Fields Applied to Low-Dimensional Structures

We want to discuss here some magneto-optical phenomena in semiconductor structures of reduced dimensionality.

At first sight many properties of low-dimensional excitons are similar to bulk samples, e.g., there is a diamagnetic shift proportional to the area of the exciton in the quantum well if $\boldsymbol{B}$ is normal to it [06G1], there is a Zeeman splitting, and there is Landau quantization. See, e.g., [91C1, 92B1, 91S1, 95B1, 99B1, 01K1, 01L1] for non-magnetic quantum structures and e.g. [96K1, 04H1, 05B1] for DMS based ones.

The development of the fan of Landau-levels in a magnetic field is nicely demonstrated for free carriers a GaAs quantum-well sample in Fig. 24.8. The heavy-hole and light-hole excitons (the features at low energies) are, how ever, not much affected by the magnetic field.

The motion dependent magnetic properties of excitons e.g. in CdTe have been investigated in [08S1, 10D1, 10L1]. The idea was to investigate relatively thick CdTe QW with a thickness larger than the excitonic Bohr radius. The relative motion of the electron–hole pair is then hardly modified, but the centre-of-mass momentum $\hbar K$ is quantized perpendicular to the plane of the QW to integer multiples N of $\hbar\pi/l_z$. Investigating the influence of the $\boldsymbol{B}$-field as a function of N allows deducing

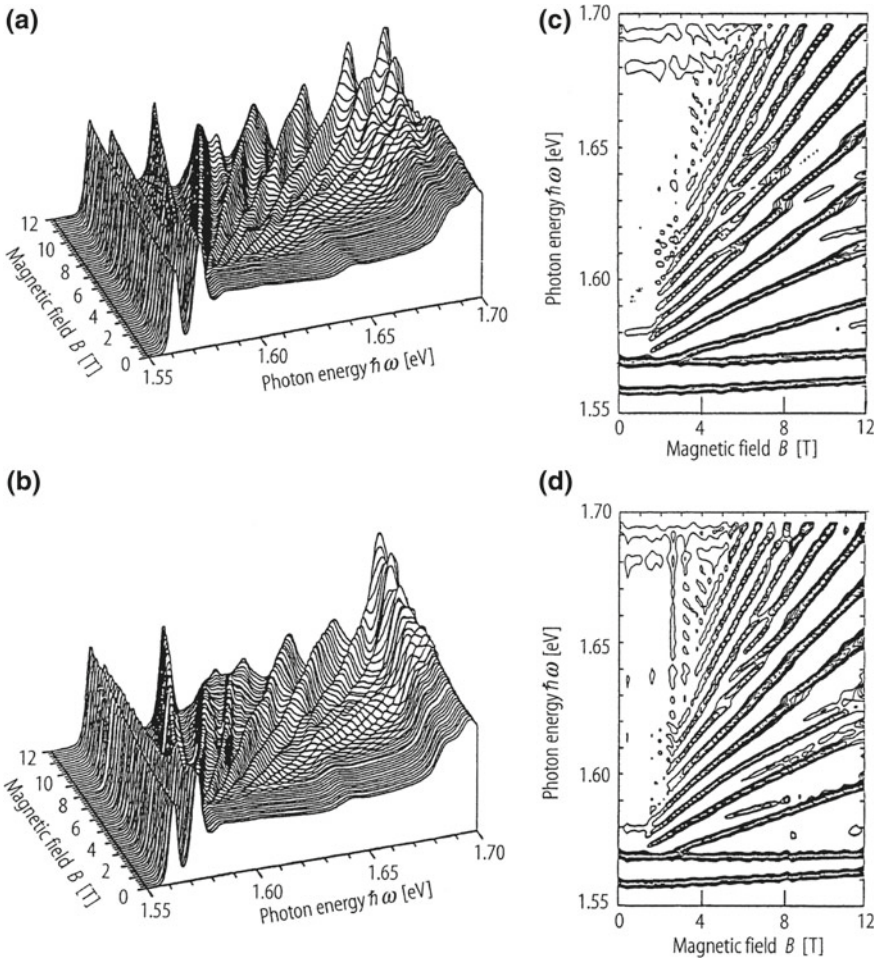


Fig. 24.8 Linear magneto absorption spectra of a GaAs quantum-well sample with 8.5 nm-wide wells for two different circular polarizations of the light (**a**, **b**). The resulting contour plots (**c**, **d**) (reprinted with permission from [91S1] ©1991 Springer Nature)

the magnetic properties in dependence of the perpendicular wave vector while the parallel one is zero for normal light incidence. Its was found, that the Zeeman effect increases with N, that the diamagnetic shift decreases with N^{-2} and the exciton effective mass increases with $\boldsymbol{B}^2$. These phenomena could be attributed to coupling between various sub-levels and are partly consequences of the magnetic Stark effect. This latter effect was also observed for GaAs/$Al_{1-x}Ga_xAs$ QW in [98L1].

A further effect in quantum wells comes from the well-width dependence of the lateral part of the wavefunctions [92I1, 95H1, 96O1, 97S1]. With decreasing well width the wave function penetrates deeper into the barrier. Consequently the

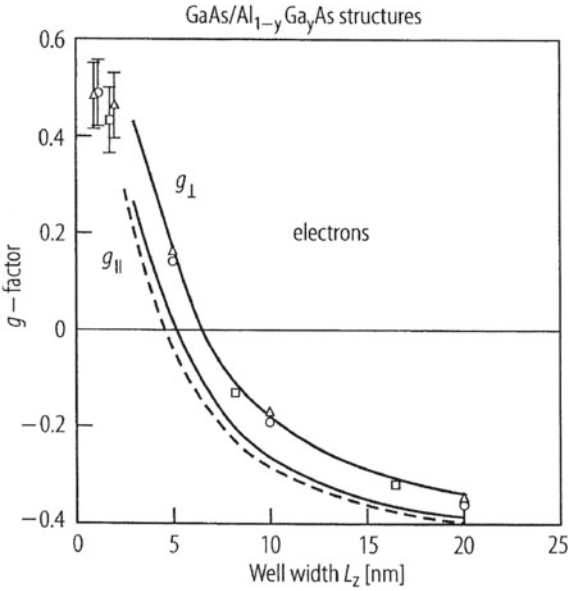


Fig. 24.9 Dependence of electron g-factors on the well width of $\boldsymbol{B}$ in $GaAs/Al_{1-y}Ga_yAs$ quantum-well structures. The symbols are experimental data from different sample while the various lines are theoretical predictions (reprinted with permission from [95H1] ©1995 Elsevier)

influence of the magnetic properties of the barrier, e.g., of its g-factors, grows. In Fig. 24.9 we show this for the example of the electron g-factor. The latter is negative in GaAs and positive in AlAs. Consequently g_e increases with decreasing well width of $GaAs/Al_{1-y}Ga_yAs$ structures and even exhibits a change of sign. Furthermore a dependence on the orientation of the magnetic field being either perpendicular or parallel to the quantum well has been theoretically predicted and observed experimentally in narrow quantum wells [97L1].

It is obvious that the strong anisotropy of low-dimensional structures leads to the fact that the magnetic field effects not only depend on the orientation of the $\boldsymbol{B}$-field to the crystallographic axes but also on its orientation relative to the quantization axes [96O1, 02B1]. The orientation of the $\boldsymbol{B}$-field perpendicular to the growth direction is called *Faraday configuration* while the one with the $\boldsymbol{B}$-field in plane is called *Voigt configuration*.

In Faraday configuration the electron and hole states in a quantum well are fully quantized in all three directions due to the magnetic field. There is still some rotational symmetry (rotations of 90°) around the $\boldsymbol{B}$-field axis. The bright excitons in GaAs/AlGaAs and InGaAs/GaAs quantum wells show a Zeeman splitting that increases linearly for small magnetic fields. The splitting gets non-linear for fields larger than about 1 T due to spin-dependent field-induced admixtures of light- and heavy-hole valence bands [97T1].

The mentioned rotational symmetry also applies to self-assembled quantum dots with negligible in-plane asymmetry (D_{2d} symmetry, see Fig. 23.29). In magnetic field (Fig. 24.10a) the bright exciton states (± 1) show (in contrast to quantum-well excitons) a splitting which increases linearly with rising field strength up to high

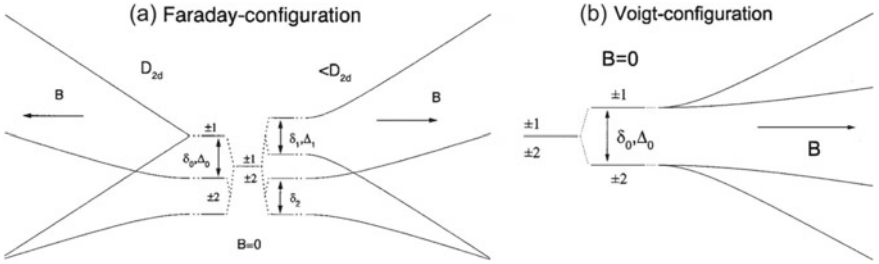


Fig. 24.10 Finestructure of quasi-0D excitons in self-assembled InGaAs/GaAs quantum dots as a function of a magnetic field $\boldsymbol{B}$ in Faraday configuration and D_{2d} (right) and $< D_{2d}$ symmetries (**a**) as well as Voigt configuration (**b**). For the notation of the various contributions to the fine splitting see text in connection with Fig. 23.29 (adapted with permission from [02B1] ©2002 American Physical Society http://doi.org/10.1103/PhysRevB.65.195315)

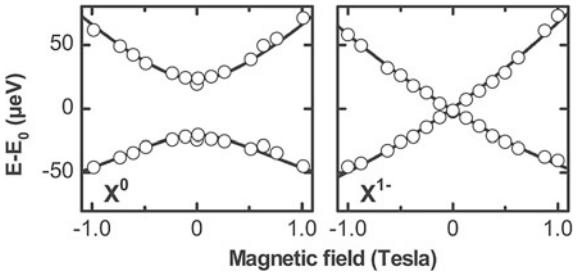


Fig. 24.11 Splitting of the bright neutral exciton (**a**) and the negative trion (**b**) in a single self-assembled InGaAs/GaAs quantum dot with $< D_{2d}$ symmetry as a function of a magnetic field $\boldsymbol{B}$ in Faraday configuration. The residual charge state of the dot can been changed since the QD layer is part of a field-effect transistor structure (reprinted with permission from [04H2] ©2004 American Physical Society http://doi.org/10.1103/PhysRevLett.93.217401)

magnetic fields. The dark excitons show a nonlinear splitting due to the hybridization of the (± 2) exciton states in the dots [02B1]. The splittings change non-linearly in all cases for reduced in-plane symmetry ($< D_{2d}$).

The Voigt configuration definitely destroys the invariance of the quantum dots with respect to in-plane rotations. The result is a mixing of bright and dark exciton states which for one makes the later appear in optical spectra. Secondly, the splittings of the exciton states are non-linear and show an anti-crossing behavior (Fig. 24.10b) [02B1].

The difference in magnetic-field behavior for neutral and charged excitons is illustrated in Fig. 24.11. The neutral bright exciton (in (a)) has a finestructure splitting (Δ_{FS}) at zero field when the dot symmetry is reduced by geometrical asymmetry or (applied) uniaxial in-plain strain [07S1] ($< D_{2d}$). The magnetic-filed dependence is hyperbolic and given by

$$\Delta E = \pm\frac{1}{2}\sqrt{\Delta_{FS}^2 + (g^*\mu_B B^2)} + \beta B^2\,, \tag{24.6}$$

where the quadratic term is due to the diamagnetic shift [04H2].

In contrast the negative trion has no exchange splitting (see Sect. 23.5). So, the zero-field splitting of the charged exciton (here X^-) is missing (Fig. 24.11b). The two sets of data points split by the magnetic field correspond to trions with the spin of the resident electron pointing upwards or downwards, respectively. The fits to the data are based on (24.6) with $\Delta_{FS} = 0$ [04H2].

The two exciton resonances in Fig. 24.11a get asymptotically circularly polarized at high magnetic fields [07S1]. The two resonances of the trion in (b) are circularly polarized with opposite orientation. Actually, when the residual electron has already a well-defined spin orientation, only the exciton transition with the electron in the respective opposite alignment is allowed. The other circular polarized exciton transition is then Pauli blocked in absorption [05H1].

For the properties of multiply charged excitons in quantum dots see [03U1, 07E1].

The above experiments demonstrate that carriers or excitons in quantum dots can be generated in well-defined spin states. The related capability of spin storage and manipulation, the dynamics of spin-relaxation and the application of the spin as qubit in spin-optoelectronics will be discussed in detail in the second volume of this textbook. For an introductory review see also [13W1].

We want to close with some remarks on low-dimensional structures based on diluted magnetic semiconductors. In quantum wells the magnetic ions can be incorporated into the barriers or into the well itself. The result is a huge Zeeman splitting and the formation of bound or free magnetic polarons [92G1, 92Y1, 96Y1]. Due to the huge Zeeman splitting it might even be possible to change the band alignment from type I to type II by a magnetic field if the respective band offset is small [86B1, 91J1, 94F1, 01L1, 03C1].

In quantum dots containing single magnetic Mn^{2+} ions strong exchange coupling to confined excitons is observed. Embedded in II–VI quantum dots like CdTe/ZnTe or CdTe/CdMgTe QDs the Mn^{2+} spin of $S = 5/2$ leads to a splitting of the exciton photoluminescence into $2S+1 = 6$ lines reflecting the Mn^{2+} (randomly fluctuating) spin orientation when the exciton recombines [04B1, 14B1]. In InAs QDs the Mn^{2+} ion is a neutral acceptor with effective spin $J = 1$ that couples ferro-magnetically to the hole but essentially not at all to the electron [07K2].

24.2 Excitons and Carriers in Electric Fields

In the last section we saw the analogy between effects of an applied $\boldsymbol{B}$-field to atoms and the quadratic (diamagnetic) and linear (Zeeman) effects experienced by excitons in semiconductors. Similar analogies can be drawn for excitons in a static electric field $\boldsymbol{E}_s$ (the ac or optical Stark effect will be covered in the second volume of this textbook). One observes a *Stark effect*, i.e., a shift (and splitting) of the exciton resonances, usually quadratic but possibly also linear in $\boldsymbol{E}_s$.

The influence of $\boldsymbol{E}_s$ on the band-to-band transitions, i.e., on the continuum states is also known as the *Franz–Keldysh effect* [61T1, 71P1]. The tilting of the electronic

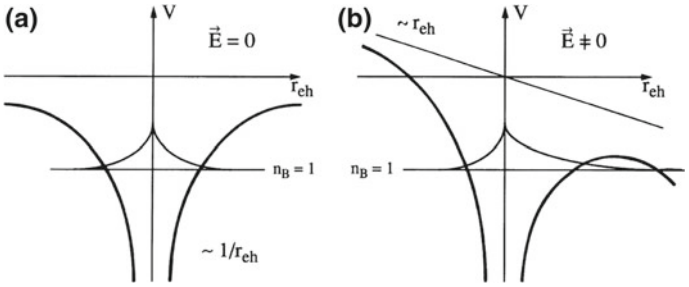


Fig. 24.12 The Coulomb potential of a three-dimensional exciton without (**a**) and with (**b**) an applied constant electric field

bands by an applied static electric field $\boldsymbol{E}_s$ allows, in this case, the Bloch-type electron and hole wave functions to tunnel over some limited distance into the forbidden gap. This results in a red-shift of the band-to-band absorption tail and the appearance of an oscillatory structure above, i.e., in the continuum states. The modifications of the absorption spectra due to $\boldsymbol{E}_s$ also influence the refractive index in the transparent spectral region below the exciton resonances according to the Kramers–Kronig relations (see Appendix A), resulting in Pockels- or Kerr-effect-like phenomena, which describe the modification of an existing or the introduction of a new birefringence, depending linearly or quadratically on $\boldsymbol{E}_s$, respectively [71M1]. We shall concentrate here on the influence of $\boldsymbol{E}_s$ on the eigenfrequencies.

24.2.1 Electric Fields Applied to Bulk Semiconductors

The observation of the Stark effect for excitons in three-dimensional (or bulk) samples is not easy, for the following reason: To get an observable shift of the eigenstates, the electric field energy $e\boldsymbol{E}_s a_B$ should be comparable with the spectral width of the absorption bands, i.e. at low temperatures $e\boldsymbol{E}_s a_B \gtrsim \Delta_{LT}$. This condition requires fields of the order of 10^6 Vm^{-1} depending on the material parameters. On the other hand, such fields broaden or even destroy the exciton resonances due to two effects. One is *field ionization* of the exciton, by tunneling through the finite Coulomb barrier at finite fields, as illustrated in Fig. 24.12. The other is *impact ionization*; this means that carriers which are always present in a semiconductor at finite temperature can gain such high energies in the electric field, that they can ionize an exciton if they hit it [90J1], resulting in two more carriers and a collision-broadening of the exciton resonance.

In Fig. 24.13 we show an example for the quadratic Stark effect of the $n_B = 1$ exciton in CdS. The shift and a slight broadening of the excitonic absorption band are clearly visible.

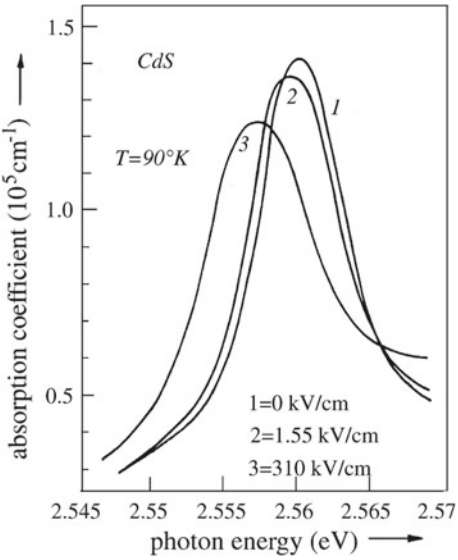


Fig. 24.13 Influence of an external field on the absorption spectrum of the B $n_B = 1$ exciton of CdS for various applied electric fields (reprinted with permission from [71L1] ©2006 John Wiley and Sons)

As mentioned above, the Franz–Keldysh oscillations appear in the band-to-band transition region. The energies of the alternating maxima and minima E_n follow the relation [04G1]:

$$(E_n - E_g)^{3/2} \sim n\,. \tag{24.7}$$

Fig. 24.14 shows modulation spectra (see Experimental Technique 16.2) of GaAs samples at 100 K exhibiting the Franz–Keldysh effect and the verification of (24.7). The Franz–Keldysh effect has here been induced by a surface electric field.

In Fig. 24.15 we give results of a rather unusual type of excitation spectroscopy. A GaSe sample is illuminated with monochromatic light through a semitransparent Au electrode and the photovoltage that develops between the contacts is measured as a function of the photon energy. This type of *photovoltage excitation spectroscopy* shows at low temperature the continuum transitions superimposed by some oscillatory structure, possibly Franz–Keldysh oscillations, which may result from the built-in field of the Schottky contact. With increasing temperatures these oscillations disappear and the exciton resonance shows up, shifts with the band gap and is washed out due to increasing homogeneous broadening. See, e.g., Figs. 21.8 and 21.10 for comparison. The exciton resonance is probably suppressed at low temperatures by field ionization in the Schottky barrier. Which features dominate depends on the various length scales like the extension of and field strength in the Schottky barrier or the penetration depth of the light.

For the above reason, not too many successful attempts have been reported to observe the Stark effect for excitons in bulk materials directly. More examples for such observations in CdS, CdSe or CuCl are found in [61T1, 69H1, 70M1, 73R1, 90J1]. Further examples for the Franz–Keldysh effect in bulk semiconductors can

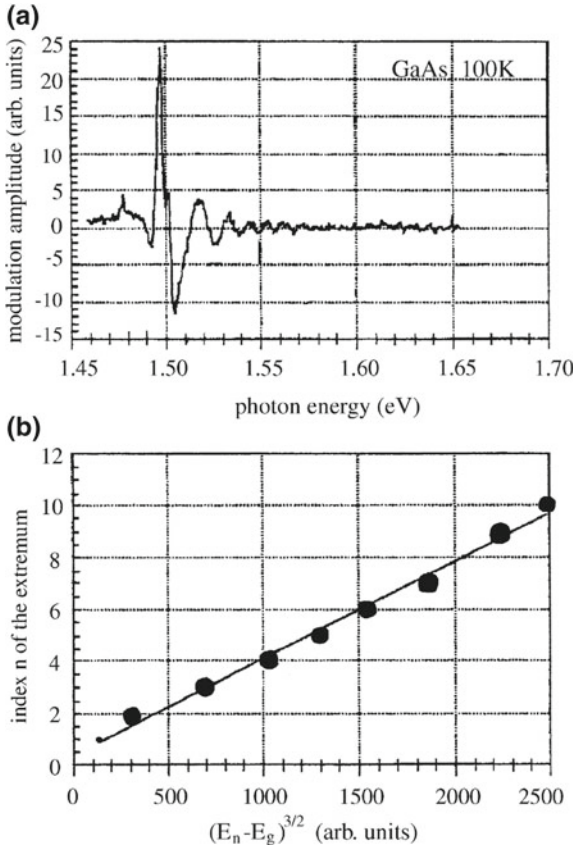


Fig. 24.14 Franz–Keldysh oscillations observed in a GaAs sample (**a**) and the relation (24.7) (**b**) [95L1]

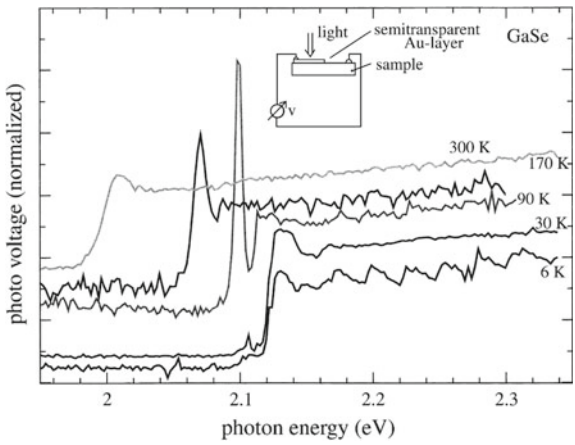


Fig. 24.15 Photovoltage excitation spectra of GaSe. The spectra are normalized in the continuum regime and shifted for clarity [02G1]

be found, e.g., in [73B1, 74Y1, 76B1] and for the Pockels effect, e.g., in ZnO in [70M1, 71M1].

Rather than applying large electric fields to observe the effects described so far one exploits the differential modifications of the electronic states occurring at low fields in the highly sensitive methods of modulation spectroscopy. We have already introduced electro-reflectance and electro-absorption in 16.2 and showed another spectrum in Fig. 24.14. For more examples refer to [67G1, 69C1, 70H1, 71M1, 73A1, 73S1, 76H1, 82L1, 84S1, 85S1, 15H1, 15K1, 18R1].

24.2.2 Electric Fields Applied to Low-Dimensional Structures

The field ionisation of excitons occurring in bulk semiconductors can be overcome by confining the electron and hole between barriers, e.g., in quantum wells, wires or dots, as shown schematically in Fig. 24.16. With increasing $\boldsymbol{E}_s$ oriented perpendicular to the layers the electron and hole shift into their respective sides of the confining potential, reducing their energetic separation. This results in a roughly quadratic narrowing of the hh-e transition and thus to a redshift of the exciton resonance.

The overlap between electron and hole wavefunction is thereby reduced, resulting in a decrease of the oscillator strength and a decrease of the exciton binding energy. This latter effect, however, is only a small correction to the reduction of the energetic separation of the first quantized levels. In addition the selection rule $\Delta n_z = 0$ is relaxed because $\boldsymbol{E}_s$ mixes states with odd and even parity, inducing some transitions which are forbidden for $\boldsymbol{E}_s = 0$. For a detailed elaboration of this so-called *quantum-confined Stark effect (QCSE)* see [85M1, 86M1] and for its application as light modulator [94E1].

The three above-mentioned effects, namely the redshift of the exciton, the decrease of its oscillator strength, and the appearance of forbidden transitions, are illustrated in Fig. 24.17, where the quantum-confined Stark effect is shown for an $In_{1-y}Ga_yAs/InP$

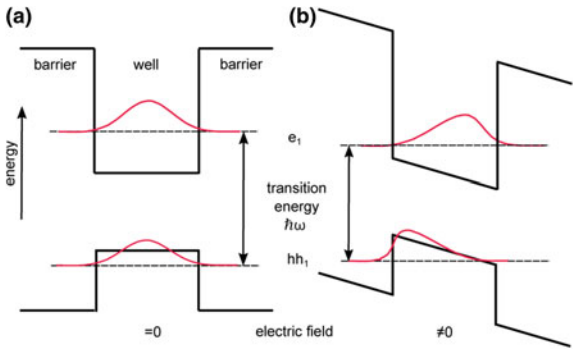


Fig. 24.16 The bandstructure of a quantum well without (**a**) and with (**b**) an applied constant electric field perpendicular to the wells, but neglecting the Coulomb interaction between electrons and holes

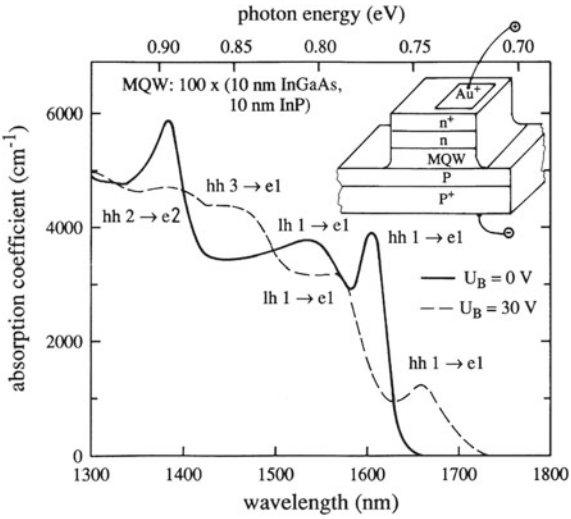


Fig. 24.17 The quantum-confined Stark effect in a InGaAs MQW sample (reprinted with permission from [87B1] ©1987 AIP Publishing)

MQW located in the intrinsic region of a reverse-biased pin diode as shown in the inset.

For the quantum confined Stark effect in $Si_{1-y}C_y$/Si quantum wells see, e.g., [99F1].

If an electron moves in a quantum well in an electric field normal to the well, the spin of the electron precesses due to spin-orbit coupling. This effect allows the manipulation of the electron spin as shown in [03R1] and discussed in detail in the second volume.

The application of an electric field normal to the planes of superlattices shifts adjacent wells energetically with respect to each other by an amount

$$\Delta E = e\,|\boldsymbol{E}_\mathrm{s}| \cdot d \tag{24.8}$$

where $\boldsymbol{E}_\mathrm{s}$ is the applied field and d is the superlattice period. This fact eventually destroys the origin of miniband formation namely the overlap of the levels of adjacent wells with equal energy. Instead a so-called *Wannier–Stark ladder* is formed. We will come back to this phenomenon in the second volume when describing Bloch oscillations. Further examples can be found, e.g., in [93M1, 94Y1, 01L1].

The quantum confined Stark and Franz–Keldysh effects are also known for structures of lower dimension like quantum wires [99W1], for quantum dots in insulating matrices [90E1] or for self-assembled quantum dots and quantum islands in QWs [88M1, 98H1, 00F1, 02W1].

It should finally be noted that the electric field of charges trapped in the vicinity of a quantum dot or island is also (at least partly) responsible for the fluorescence intermittency of single dots (see, e.g., [01K2]).

For a review of $\boldsymbol{E}$-field induced effects in quantum structures see e.g. [91E1].

24.3 Influence of Strain Fields on Electronic and Excitonic States

The third external or internal perturbation which we shall consider here are mechanical strain fields. The crucial quantity which enters here is the dependence of the energy of the band extrema on the strain, the so-called *deformation potential*. It is defined for conduction and valance band by (see (14.24))

$$\Xi = a\frac{\mathrm{d}E_{\mathrm{c,v}}}{\mathrm{d}x} \tag{24.9}$$

where a is the lattice constant, $\mathrm{d}E_{\mathrm{c,v}}$ is the shift of the band edge caused by a length variation of the unit cell by an amount $\mathrm{d}x$. Deformation potentials are defined for various types of strain like hydrostatic, uniaxial or shear strain. For details see, e.g., [82L2, 96Y1] or Sect. 14.6.

For excitons the relevant quantity is the sum of the deformation potentials of conduction and valence bands. As in the case of magnetic and electric fields, an applied stress changes the eigenenergies of the exciton states and may also lead to a splitting of degenerate states, if the resulting strain is oriented such that it reduces the symmetry of the crystal. The latter situation is, e.g., realized when stress is applied perpendicular to the crystallographic $\boldsymbol{c}$ axis of a uniaxial material (e.g., the Wurtzite structure) where it lifts the degeneracy in the plane perpendicular to $\boldsymbol{c}$, but not if the stress acts in the direction of $\boldsymbol{c}$. This orientation leads only to a shift of eigenenergies but not to a splitting since the symmetry is not changed. For details see Appendix B on group theory.

24.3.1 Strain Fields in Bulk Semiconductors

We now give various examples for the impact of strain fields. The shift, the splitting and change in oscillator strength of excitonic transition subject to *uniaxial strain* are demonstrated in Fig. 24.18. We use the example of the reflection spectra of CdS in the range of the $n_{\mathrm{B}} = 1\mathrm{A}\Gamma_5$ and $\mathrm{B}\Gamma_5$ exciton resonances for zero and finite stress perpendicular to the $\boldsymbol{c}$ axis. Measurements of this type can be used to determine the deformation potentials. For materials with zinc-blende and Wurtzite type structures see [62S1, 70L1, 72G1, 79R1, 79T1, 91R1, 96G1].

Sometimes, the strain is induced during growth of layers, e.g., on non-lattice matched substrates. Examples for such strained layers can be found in [89O1]. An example for growth induced strain in ZnO is [99R1]. Note in this context the comments on the valence band structure in [10K1]. A further contribution to strain may come from different thermal expansion coefficients of substrate and layer. Mechanisms of strain relaxation are discussed in [12L1] and references therein.

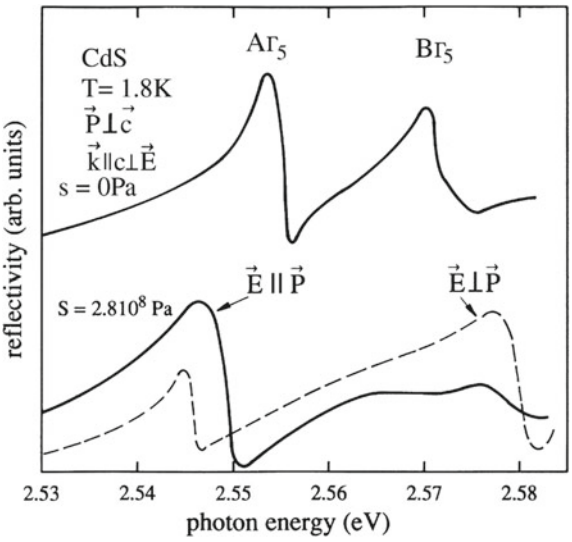


Fig. 24.18 The excitonic reflection spectra of CdS under uniaxial stress (based on data from [70L1])

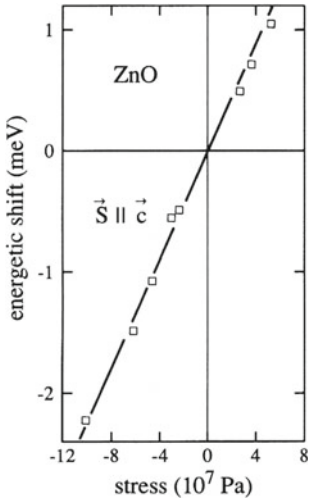


Fig. 24.19 The shift of a BEC luminescence line in ZnO under the influence of strain (reprinted with permission from [68S1] ©1968 Springer Nature)

Since bound exciton complexes (BEC) often have very narrow absorption and emission bands (Sect. 22.1) one can easily study the influence of magnetic or strain fields, which affect both the ground and the excited states as already mentioned above. In Fig. 24.19 we give the shift of the luminescence of a BEC in ZnO which shows clearly the influence of the strain. According to arguments given above, a shift but no splitting occurs for the orientation $\boldsymbol{S} \parallel \boldsymbol{c}$ in contrast to the situation of Fig. 24.17. For more details see [68S1].

There are two contributions to *biaxial strain*, which occurs in the plane of epitaxial growth, and which can be considered as a superposition of a hydrostatic and a uniaxial

strain. One contribution to the strain already arises during the growth process if the lattice constants of substrate and layer do not coincide at growth temperature. The first epitaxial layers will often grow with some strain, trying to match the lattice constant of the substrate. With increasing layer thickness it is energetically more favorable for the layer to create a dislocation network which relaxes the strain [85P1, 87D1], or to form islands. This latter growth mode is known as Stranski–Krastanov mode (see Appendix D). Sometimes a small residual strain remains, independent of the layer thickness [92G1, 92W1, 92W2]. The critical thickness l_c for the onset of dislocation formation depends on the lattice misfit, on the elastic constants of the epilayer, on the energy necessary to produce the dislocations, on the surface energies etc. It can be as small as a few atomic layers only. The dislocation network explains the function of a buffer layer or a short period superlattice, which are often used if two materials of different lattice constant are grown on top of each other. In this case the crystalline quality of the interface is considerably poorer than that of the surface of the layer after a thickness of about 0.1 μm.

The other part of the biaxial strain arises during the cooling from the growth temperature to the temperature at which the (optical) measurements are performed if the coefficients of thermal expansion are different for substrate and layer, as is usually the case. A part of this thermal strain can sometimes be accommodated by modifications in the dislocation network. In Fig. 24.20 we show reflection spectra in the region of the $n_B = A\Gamma_5$ and $B\Gamma_5$ excitons in CdS grown on SrF_2.

Since the cubic SrF_2 substrate was oriented (111), the hexagonal CdS grows with the $\boldsymbol{c}$ axis perpendicular to the interface. Consequently the C_{6v} symmetry is not affected by the biaxial strain, but the eigenenergies are shifted. The bulk values for the longitudinal eigenenergies (or the reflection minima) are given by the vertical lines. In agreement with the discussion above, we see that the excitons are shifted and

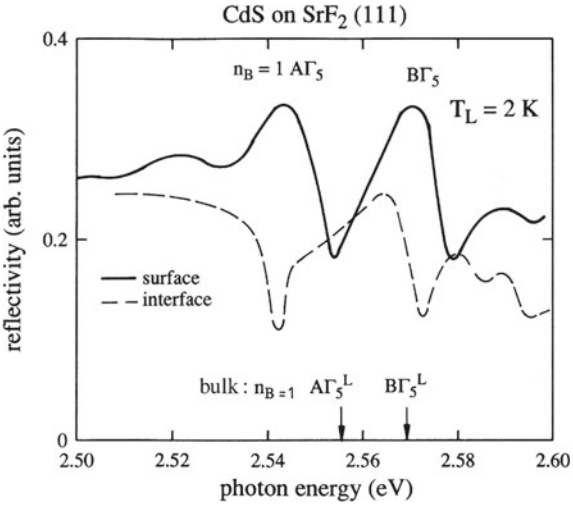


Fig. 24.20 Reflection spectra of the $n_B = 1A\Gamma_5$ and $B\Gamma_5$ excitons of a CdS layer grown on SrF_2 and measured at the surface and at the interface. The positions of the longitudinal eigenenergies in the bulk are indicated by vertical arrows [91G1]

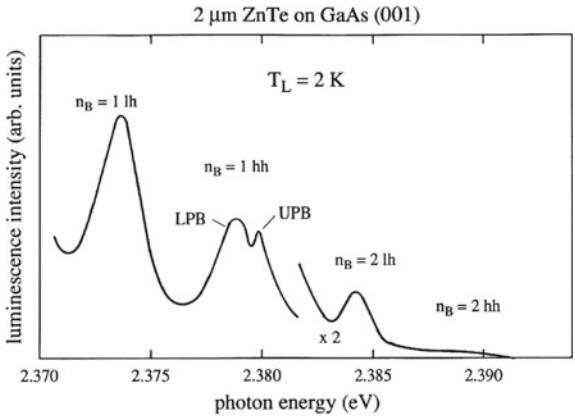


Fig. 24.21 Luminescence spectrum of a ZnTe layer grown on GaAs (001), showing the strain-induced splitting of the light- and heavy-hole $n_B = 1$ exciton resonances (reprinted with permission from [92W1] ©1992 Springer Nature). Compare with Fig. 24.4a

that this shift is different at the interface and the surface. Furthermore, the damping is larger at the interface than at the surface due to the dislocations in the former region.

In contrast to the situation in Fig. 24.20, the luminescence spectrum of Fig. 24.21 was taken on the cubic material ZnTe grown on GaAs(00l). In this case, the biaxial strain reduces the cubic T_d symmetry of bulk ZnTe and the Γ_5 exciton resonance seen in Fig. 21.4 splits into its light and heavy hole components [92G1, 92W1, 92W2].

The biaxial in-plane tensile strain results in this case essentially from the different thermal expansion coefficients of the GaAs and ZnTe layers since the ZnTe layer thickness of 2 μm is well above l_c and the ZnTe layer grows with its own lattice constant.

Inhomogeneous strain can actually be used to create potential traps for electron-hole pairs. For examples see, e.g., [89O1], the discussion of electron–hole liquids in Si or Ge given in the second volume, and [03N1] for the case of Cu_2O. However, strain resulting from the fixation of the sample on the holder in a cryostat is usually unwanted and should be avoided.

24.3.2 Strain in Structures of Reduced Dimensionality

The influence of strain and of strain relaxation is of course very important when growing epilayers, quantum wells, superlattices or self-assembled quantum dots on non-lattice matched substrates (see, e.g., [10K1, 12L1]).

The above-mentioned strain-induced splitting of the Γ_8 valence band is of special importance for strained quantum wells and superlattices whenever a layer is grown with a lattice constant different from that of the substrate. The splitting adds to the one resulting from the different effective masses of the lh and hh bands entering in to the confinement energies.

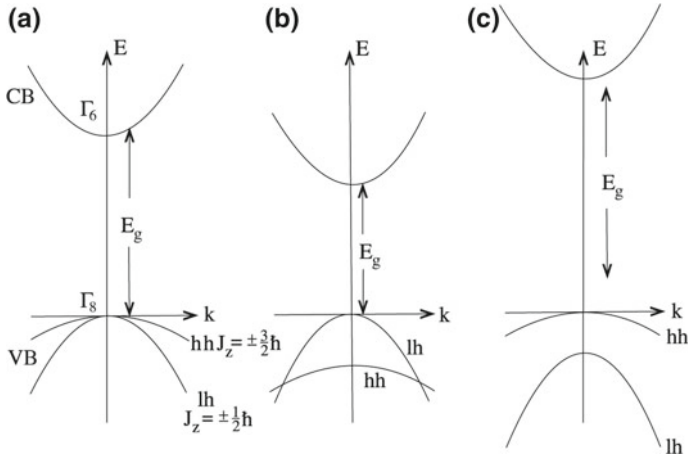


Fig. 24.22 Schematic drawing of the influence of strain on the alignment of the hh and lh valence bands and on the value of the gap for zero strain (**a**), for two-dimensional tensile strain and normal uniaxial compressive strain (**b**), and for two-dimensional compressive strain and normal uniaxial tensile strain (**c**). See also Fig. 14.9

A two-dimensional tensile strain in the plain of the layer is usually accompanied by a uniaxial compressive strain normal to it and vice versa. The results of this strain on the width of the gap and on the arrangement of the hh and lh valence bands is shown schematically in Fig. 24.22 and verified for tensile strain, e.g., in Fig. 24.20.

In quantum wells and superlattices (and in structures of even lower quasi-dimensionalities) the strain induced effects come in addition to the quantization energies, which are always larger for the lh exciton compared to the hh exciton. Consequently, strain effects may enhance or reduce the splitting caused by confinement. For strained superlattices one frequently uses a substrate with a lattice constant situated between the values of the two different layer materials of the superlattice resulting in so-called strain-symmetrized superlattices because the layers show alternating compressive and tensile strain. A similar situation can be reached in *free standing superlattices*, which occur when the thickness of the superlattice is so thick that the coherence of its lattice constant to the substrate is lost, e.g., by the formation of dislocations.

While many III–V systems can be grown lattice-matched to GaAs, InAs or InP substrates, the II–VI and IV–IV systems are usually strained, as seen in Fig. D.2. Typical examples are, e.g., $Si_{1-y}Ge_y$/Si superlattices (e.g., [97K1] and references above), CdS/ZnS [97H1] or CdS/ZnSe [99P1, 01D1]. For the influence of pressure on symmetric and asymmetric short period superlattices see e.g. [90H1, 91N1, 92R1].

Further important implications of strain are the following: The strain in superlattices results via the piezo effect also in alternating electric fields. For such piezo-superlattices see Sect. 23.2 or [99A1]. For the creation of an inhomogeneous strain, acting as a potential trap for excitons in quantum wells see, e.g., [99N1]. The

finestructure splitting of excitons in quantum dots due to an asymmetric in-plane potential (anisotropic exchange splitting) can be modified by the application of in-plane uniaxial strain [06S1].

Problems

24.1. Typical values of the deformation potential are around 10 eV. What is the shift of the band edges for $\Delta a/a = 10^{-3}$?

24.2. Do you expect the band-gap shifts due to the lattice deformation by acoustic and by optical phonons to be identical or not? Why?

24.3. Calculate the diamagnetic shift at $\boldsymbol{B} = 10$ T for excitons characterized by $E_{\mathrm{ex}}^{\mathrm{b}} = 5\,\mathrm{meV}$, $\varepsilon = 15$ and by $E_{\mathrm{ex}}^{\mathrm{b}} = 100\,\mathrm{meV}$, $\varepsilon = 6$, respectively.

24.4. Calculate the Zeeman splitting for a spin singlet and a triplet exciton for $n_{\mathrm{B}} = 1$, $g_{\mathrm{e}} = 1.6$ and $g_{\mathrm{h}} = 2.2$. What is the difference if at $\boldsymbol{B} = 0$ the two states already show a finite energy splitting δ?

24.5. Do you expect that the linear Stark effect can occur for $n_{\mathrm{B}} = 1$ and/or for $n_{\mathrm{B}} = 2$ excitons? Why?

24.6. Calculate in the simplest approximation the Stark shift for excitons with the data given in connection with Problem 3 and for an electric field strength E of $10^3\ \mathrm{Vm}^{-1}$, $10^6\ \mathrm{Vm}^{-1}$, and $10^{-2}\ E_{\mathrm{ex}}^{\mathrm{b}}(ea_{\mathrm{B}})^{-1}$.

24.7. Calculate the quantum confined Stark effect by perturbation theory for a quantum well with infinitely high barriers and the other data as for the hh exciton in GaAs or InAs. Compare with experimental data.

24.8. Proof whether the two examples given in the introduction to this chapter are correct.

References

[60H1] J.J. Hopfield, J. Phys. Chem. Solid **15**, 97 (1960)
[61T1] D.G. Thomas, J.J. Hopfield, Phys. Rev. **124**, 657 (1961)
[62S1] M.D. Sturge, Phys. Rev. **127**, 768 (1962)
[67G1] E. Gutsche, H. Lange, Phys. Status Solidi (b) **22**, 229 (1967)
[68S1] Chr. Solbrig, Z. Phys. A **211**, 429 (1968)
[69C1] M. Cardona, *Modulation Spectroscopy* (Academic, New York, 1969)
[69H1] C. Gahwiller, G. Harbeke, Phys. Rev. **185**, 1147 (1969)
[70H1] N. Hase, M. Onuka, J. Phys. Soc. Jpn. **28**, 965 (1970)
[70L1] D.W. Langer et al., Phys. Rev. B **2**, 4005 (1970)
[70M1] E. Mohler, Phys. Status Solidi **38**, 81 (1970)

[71L1] H. Lange, Phys. Status Solidi (b) **48**, 791 (1971)
[71M1] O.W. Madelung, E. Mollow, Z. Phys. A **249**, 12 (1971)
[71P1] J.I. Pankove, *Optical Processes in Semiconductors* (Dover Publications, New York, 1971)
[72G1] O. Goede et al., Phys. Status Solidi B **53**, 235 (1972)
[73A1] D.E. Aspnes, A.A. Studna, Phys. Rev. B **7**, 4605 (1973)
[73B1] B.A. Bobylev, A.F. Kravchenko, A.S. Terekhov, Sov. Phys. Semicond. **6**, 1635 (1973)
[73H1] K. Hümmer, Phys. Status Solidi (b) **56**, 249 (1973)
[73R1] B.S. Razbirin, I.N. Uraltsev, A.A. Bogdanov, Sov. Phys. Solid State **15**, 604 (1973)
[73S1] B.O. Seraphin (ed.), *Modulation Spectroscopy* (North-Holland, Amsterdam, 1973)
[73S2] S. Suga, T. Koda, T. Mitani, Phys. Status Solidi b **55**, 355 (1973)
[74F1] J.U. Fischbach et al., Phys. Status Solidi B **66**, 253 (1974)
[74O1] G. Ottaviani et al., Solid State Commun. **14**, 933 (1974)
[74Y1] B.G. Yacobi, Y. Brada, Phys. Rev. B **10**, 665 (1974)
[75C1] K. Cho et al., Phys. Rev. B **11**, 1512 (1975); ibid. B **12**, 1608 (1975)
[76B1] W. Buckel, U. Birkholz, K. Ziegler, Phys. Status Solidi (b) **78**, K23 (1976)
[76H1] Y. Hamakawa, T. Nishino, in *Optical Properties of Solids, New Developments*, ed. by B.O. Seraphim (North-Holland, Amsterdam, 1976), p. 255
[77V1] H. Venghaus, S. Suga, K. Cho, Phys. Rev. B **16**, 4419 (1977)
[78G1] J.A. Gaj et al., Solid State Commun. **25**, 193 (1978); ibid. **29**, 435 (1979)
[78R1] B.S. Razbirin et al., JETP Lett. **27**, 341 (1978)
[79E1] K. Cho (ed.), *Excitons*. Topics in Current Physics, vol. 14 (Springer, Berlin, 1979)
[79R1] R. Ranvaud et al., Phys. Rev. B **20**, 701 (1979)
[79T1] H.-R. Trebin, U. Rössler, R. Ranvaud, Phys. Rev. B **20**, 686 (1979)
[80S1] S. Suga et al., Phys. Rev. B **22**, 4931 (1980)
[81B1] G. Blattner et al., Phys. Status Solidi (b) **107**, 105 (1981)
[82B1] G. Blattner et al., Phys. Rev. B **25**, 7413 (1982)
[82H1] H.W. Hölscher, A. Nöthe, Ch. Uihlein, Phys. B+C **117/118**, 2379 (1982)
[82K1] C. Klingshirn, in *Application of High Magnetic Fields in Semiconductor Physics*, ed. by G. Landwehr. Lecture Notes in Physics, vol. 177 (Springer, Berlin, 1982), p. 214
[82L1] G. Landwehr (ed.), *Application of High Magnetic Fields in Semiconductor Physics*. Lecture Notes in Physics, vol. 177 (Springer, Berlin, 1982)
[82L2] O. Madelung, U. Rössler (ed.), *Landolt–Börnstein*. New Series, Group III, vol. 17a to i, 22a and b, 41A to D (Springer, Berlin, 1982–2001)
[84L1] Y.R. Lee, A.K. Romdas, Solid State Commun. **51**, 861 (1984)
[84S1] L. Schultheis, J. Lagois, Phys. Rev. B **29**, 6784 (1984)
[85H1] B. Hönerlage et al., Phys. Rep. **124**, 161 (1985)
[85M1] D.A.B. Miller et al., Phys. Rev. B **32**, 1043 (1985)
[85P1] R. People, J.C. Bean, Appl. Phys. Lett. **47**, 322 (1985)
[85S1] L. Schultheis, C.W. Tu, Phys. Rev. B **32**, 6978 (1985)
[86B1] J.A. Brum, G. Bastard, M. Voos, Solid State Commun. **59**, 561 (1986)
[86B2] F. Beerwerth, D. Fröhlich, Phys. Rev. Lett. **57**, 1344 (1986)
[86G1] J. Gutowski, Solid State Commun. **58**, 523 (1986)
[86M1] D.A.B. Miller, D.S. Chemla, S. Schmitt-Rink, Phys. Rev. B **33**, 6976 (1986)
[87B1] I. Bar-Joseph et al., Appl. Phys. Lett. **50**, 1010 (1987)
[87D1] B.W. Dodson, J.Y. Tsao, Appl. Phys. Lett. **51**, 1325 (1987)
[88G1] J.A. Gaj, in *Semiconductors and Semimetals*, vol. 25, ed. by J.K. Furdyna, J. Kossut (Academic, New York, 1988), p. 275
[88G2] O. Goede, W. Heimbrodt, Phys. Status Solidi B **146**, 11 (1988)
[88M1] D.A.B. Miller, D.S. Chemla, S. Schmitt-Rink, Appl. Phys. Lett. **52**, 2154 (1988)
[88S1] J.K. Furdyna, J. Kossut (eds.), *Semiconductors and Semimetals*, vol. 25. Growth and Optical Properties of Wide Gap II-VI Low Dimensional Semiconductors (Academic, New York, 1988)
[89G1] J. Gutowski, NATO ASI Series B, vol. 200 (Plenum, New York, 1989), p. 139
[89O1] O'Reilly, Semicond. Sci. Technol. **4**, 121 (1989)

[90E1] A.I. Ekimov et al., J. Lumin. **46**, 83 (1990)
[90H1] M. Holtz et al., Phys. Rev. B **41**, 3641 (1990)
[90J1] M.A. Jackobson et al., Phys. Status Solidi (b) **161**, 395 (1990)
[91C1] D.S. Chemla, in *Highlights in Condensed Matter Physics and Future Prospects*, ed. by L. Esaki (Plenum Press, New York, 1991), p. 293
[91E1] L. Esaki, in *Highlights in Condensed Matter Physics and Future Prospects*, ed. by L. Esaki (Plenum Press, New York, 1991), p. 55
[91G1] H. Gießen, Diploma thesis, Universität Kaiserslautern (1991)
[91J1] B.T. Jonker et al., J. Appl. Phys. **69**, 6097 (1991)
[91N1] J. Nunnenkamp et al., Phys. Rev. B **44**, 8129 (1991)
[91O1] M.v. Ortenberg, Festkörperprobleme/Adv. Solid State Phys. **31**, 261 (1991)
[91R1] K. Reimann et al., Phys. Rev. B **44**, 2985 (1991)
[91S1] S. Schmitt-Rink, Festkörperprobleme/Adv. Solid State Phys. **31**, 243 (1992)
[92B1] L.V. Butov, Phys. Rev. B **46**, 12765, 13627 and 15156 (1992)
[92G1] W. Gebhardt, Mater. Sci. Eng. B **11**, 1 (1992)
[92G2] O. Goede, W. Heimbrodt, Festkörperprobleme/Adv. Solid State Phys. **32**, 237 (1992)
[92I1] E.L. Ivchenko, A.A. Kiselev, Sov. Phys. Semicond. **26**, 827 (1992)
[92R1] T. Rappen, G. Mohs, M. Wegener, Phys. Status Solidi (b) **173**, 77 (1992)
[92W1] H.P. Wagner, H. Leiderer, Festkörperprobleme/Adv. Solid State Phys. **32**, 221 (1992)
[92W2] H.P. Wagner et al., J. Lumin. **52**, 41 (1992)
[92Y1] D.R. Yakovlev, Festkörperprobleme/Adv. Solid State Phys. **32**, 237 (1992)
[93M1] E.E. Mendez, in *Optics of Semiconductor Nanostructures*, ed. by F. Henneberger, S. Schmitt-Rink, E.O. Göbel (Academic, Berlin, 1993), p. 181
[94E1] U. Efron, G. Livescu, in *Spatial Light Modulator Mater Devices and Technology*, ed. by U. Efron (M. Dekker Inc., New York, 1994), p. 217
[94F1] J.K. Furdyna, Solid State Electron. **37**, 1065 (1994)
[94G1] J.A. Gaj et al., Phys. Rev. B **50**, 5512 (1994)
[94Y1] M. Yamaguchi et al., Solid State Electron. **37**, 839 (1994)
[95B1] L.V. Butov et al., Phys. Rev. B **52**, 12153 (1995)
[95H1] R.M. Hannak et al., Solid State Commun. **93**, 313 (1995)
[95L1] D. Liu, MS thesis, Boston College (1995)
[95M1] H. Heinrich, J.B. Mullin (eds.), Mater. Sci. Forum **182–184** (1995)
[95S1] U. Siegner et al., Phys. Status Solidi (b) **188**, 361 (1995); Phys. Rev. B **51**, 495 (1995)
[96G1] B. Gil, F. Hamdani, H. Morkoç, Phys. Rev. B **54**, 7678 (1996)
[96K1] I.V. Kukushkin, V.B. Timofeev, Adv. Phys. **45**, 147 (1996)
[96O1] M. Oestreich, S. Hallstein, S. Rühle, IEEE J. Sel. Top. Quantum Electron. **2**, 747 (1996)
[96Y1] D.R. Yakovlev, Comment Condens. Matter Phys. **18**, 51 (1996)
[97B1] S. Bar-Ad, Phys. Rev. Lett. **78**, 1363 (1997)
[97H1] M. Hetterich et al., Phys. Rev. B **56**, 12396 (1997)
[97K1] S. Kim et al., Phys. Rev. B **55**, 7130 (1997)
[97L1] P. Le Jeune et al., Semicond. Sci. Technol. **12**, 380 (1997)
[97P1] G. Landwehr, W. Ossau (eds.), *Proceedings of 12 International Conference on High Magnetic Fields in Semiconductors*, vols. 1 and 2 (World Scientific, Singapore, 1997)
[97T1] N.J. Traynor, R.J. Warburton, M.J. Snelling, R.T. Harley, Phys. Rev. B **55**, 15701 (1997)
[97S1] A.A. Sirenko in *Proceedings of 12 International Conference on High Magnetic Fields in Semiconductors*, vol. 2, ed. by G. Landwehr, W. Ossau (World Scientific, Singapore, 1997), p. 561
[98H1] W. Heller, U. Bockelmann, G. Abstreiter, Phys. Rev. B **57**, 6270 (1998)
[98L1] S.K. Lyo, Phys. Rev. B **58**, R10187 (1998)
[98O1] H. Ohno, Science **281**, 951 (1998)
[99A1] O. Ambacher et al., Phys. Status Solidi B **216**, 381 (1999)
[99B1] M. Bayer et al., Phys. Rev. Lett. **82**, 1748 (1999)
[99F1] S. Fukatsu et al., J. Lumin. **80**, 429 (1999)

[99N1] V. Negoita, D.W. Snoke, K. Eberl, Appl. Phys. Lett. **75**, 2059 (1999); Phys. Rev. B **60**, 2661 (1999)
[99P1] S. Petillon et al., J. Cryst. Growth **201/202**, 453 (1999)
[99R1] D.C. Reynolds et al., J. Appl. Phys. **86**, 5589 (1999)
[99W1] H. Weman et al., Appl. Phys. Lett. **74**, 2334 (1999)
[00D1] T. Dietl et al., Science **287**, 1019 (2000)
[00F1] P.W. Fry et al., Phys. Rev. Lett. **84**, 733 (2000)
[01D1] A. Dinger et al., Phys. Rev. B **64**, 245310 (2001)
[01K1] R. Kotlyar et al., Phys. Rev. B **63**, 085310 (2001)
[01K2] M. Kuno et al., J. Chem. Phys. **115**, 1028 (2001)
[01L1] C. Klingshirn (ed.), *Landolt–Börnstein*. New Series, Group III, vol. 34C, Parts 1 and 2 (Springer, Berlin, 2001 and 2004); Part 3, ed. by E. Kasper, C. Klingshirn (2007)
[01O1] M. Oestreich et al., Adv. Solid State Phys. **41**, 173 (2001)
[02B1] M. Bayer et al., Phys. Rev. B **65**, 195315 (2002)
[02D1] T. Dietl, Semicond. Sci. Technol. **17**, 377 (2002)
[02G1] O. Gogolin et al., Private Communication (2002)
[02K1] K. Chang et al., Appl. Phys. Lett. **80**, 1788 (2002)
[02L1] W.L. Lambrecht et al., Phys. Rev. B **65**, 075207 (2002)
[02S1] H. Ohno (ed.), Semiconductor spintronics. Semicond. Sci. Technol. **17**(4) (2002)
[02S2] S. Sonda, IEEE Trans. Mag. **38**, 2859 (2002)
[02S3] K. Sato, H. Katayama-Yoshida, Semicond. Sci. Technol. **17**, 367 (2002)
[02W1] O. Wolst et al., Phys. E **13**, 283 (2002); ibib. E **17**, 554 (2003)
[03C1] G.P. Collins, Sci. Am. **288**, 13 (2003)
[03C2] S.F. Chichibu et al., J. Appl. Phys. **93**, 756 (2003)
[03D1] T. Dietl, Europhys. News **34**, 216 (2003); Phys. Status Solidi b **240**, 433 (2003)
[03D2] T. Dietl, H. Ohno, MRS Bull. **28**, 714 (2003)
[03O1] D. Kholkhlov (ed.), *Lead Chalcogenides Physics and Applications*. Optoelectronic Properties of Semiconductors and Superlattices, vol. 18. (Taylor and Francis, New York, 2003)
[03N1] N. Naka, N. Nagasawa, Phys. Status Solidi (b) **238**, 397 (2003)
[03R1] E.I. Rashba, A.L. Efros, Appl. Phys. Lett. **83**, 5295 (2003)
[03U1] B. Urbaszek et al., Phys. Rev. Lett. **90**, 4 (2003)
[04A1] S. Adachi, EXCON 04, Cracow, J. Lumin. **112**, 34 (2005)
[04B1] L. Besombes et al., Phys. Rev. Lett. **93**, 207403 (2004)
[04G1] G.J. Goldsmith in *Frontiers of Optical Spectroscopy (2003)*. NATO Science Series II, vol. 168 (Kluwer, Dordrecht, 2005)
[04H1] W. Heimbrodt, P.J. Klar, *Encyclopedia of Nanoscience and Nanotechnology*, vol. 4 (American Scientific Publishers, Stevenson Ranch, 2004), p. 835
[04H2] A. Högele et al., Phys. Rev. Lett. **93**, 217401 (2004)
[04M1] B. Meyer et al., Phys. Status Solidi (b) **241**, 231 (2004)
[05B1] F.J. Brieler et al., Eur. J. Inorg. Chem. **2005**, 3597 (2005)
[05D1] M.F. Doty et al., Phys. Rev. B **71**, 201201(R) (2005)
[05H1] A. Högele et al., Appl. Phys. Lett. **86**, 221905 (2005)
[05K1] K.-M.C. Fu et al., Phys. Rev. Lett. **95**, 187405 (2005)
[05O1] Ü. Özgür et al., J. Appl. Phys. Rev. **98**, 041301 (2005)
[06G1] M. Grochal, F. Grosse, R. Zimmermann, Phys. Status Solidi C **10**, 3492 (2006)
[06H1] M. Hetterich et al., Phys. Status Solidi B **243**, 812 (2006)
[06K1] C. Klingshirn et al., Phys. J. **5**(1), 33 (2006)
[06P1] J. Philip et al., Nat. Mater. **5**, 298 (2006)
[06S1] S. Seidl et al., Appl. Phys. Lett. **88**, 203113 (2006)
[07B1] J. Brandt et al., Phys. Rev. Lett. **99**, 217403 (2007)
[07D1] T. Dietl, J. Phys. Condens. Matter **19**, 165204 (2007)
[07E1] M. Ediger et al., Phys. Rev. Lett. **98**, 98 (2007)
[07K1] S. Kuroda et al., Nat. Mater. **6**, 440 (2007)
[07K2] A. Kudelski et al., Phys. Rev. Lett. **99**, 247209 (2007)

[07S1] S. Seidl et al., Phys. Status Solidi **204**, 381 (2007)
[08A1] K. Ando et al., Phys. Rev. Lett. **100**, 067204 (2008)
[08S1] L.C. Smith et al., Phys. Rev. B **78**, 085204 (2008)
[08S2] C. Sandfort et al., Phys. Rev. B **78**, 045201 (2008)
[10D1] J.J. Davies et al., Phys. Status Solidi B **247**, 1521 (2010)
[10K1] C. Klingshirn, B.K. Meyer, A. Waag, A. Hoffmann, J. Geurts, *Zinc Oxide: From Fundamental Properties Towards Novel Applications*. Springer Series in Materials Science, vol. 120 (Springer, Heidelberg, 2010)
[10L1] D.K. Loginov et al., Phys. Status Solidi B **247**, 1528 (2010)
[11C1] S.L. Chen, W.M. Chen, I.A. Buyanova, Appl. Phys. Lett. **99**, 091909 (2011)
[12L1] C. Klingshirn (ed.), *Landolt–Börnstein*. New Series, Group III, vol. 34A (Springer, Heidelberg, in preparation for 2012)
[13W1] R.J. Warburton, Nat. Mater. **12**, 483 (2013)
[14B1] L. Besombes, H. Boukari, Phys. Rev. B **89**, 085315 (2014)
[15H1] C. Huber et al., Phys. Rev. B **92**, 075201 (2015). https://doi.org/10.1103/PhysRevB.92.075201
[15K1] C. Krämmer et al., Appl. Phys. Lett. **107**, 222104 (2015). https://doi.org/10.1063/1.4936649
[18R1] F. Ruf et al., Appl. Phys. Lett. **112**, 083902 (2018). https://doi.org/10.1063/1.5017943

Chapter 25
Review of the Linear Optical Properties

In this brief chapter, we want to review and summarize some of the aspects of the linear optical properties of semiconductors that have been presented in the preceding chapters in detail.

A schematic overview of the spectra of the complex *dielectric function*, of the complex *index of refraction* and of the *reflectivity* for a typical direct-gap semiconductor covering the whole spectral range from the IR to the UV is given in Fig. 25.1. For simplicity, we consider only one optically active phonon mode, one exciton resonance, and one further resonance which represents all continuum and band-to-band transitions. We include some damping for every resonance but neglect details including spatial dispersion. The additive character of the resonances in the dielectric function is clearly visible. The background dielectric constant ε_b of one resonance is simultaneously the static one ε_s for the next higher resonance.

Figure 25.1a includes contributions which might come from two other effects. The dashed-dotted line gives the modifications introduced by a carrier *plasma*. This plasma can be caused either by high doping levels (Chap. 19) or by high optical excitation (see next volume of this book). In the first case it consists of either electrons or of holes, in the second case it is a bipolar plasma. Plasmas can also occur for high electric injection or in inversion layers at interfaces. Recalling that the transverse eigenfrequency of a plasma is zero, one immediately obtains the dashed-dotted line in Fig. 25.1.

The presence of a plasma also influences the other resonances, e.g., the phonon resonance, resulting in plasmon–phonon mixed states (Sect. 19.3), or the exciton resonances by increasing damping, by screening of the Coulomb attraction of electron and hole and by bandgap renormalization and band filling (see next volume). These effects are not shown in Fig. 25.1 for sake of clarity.

The second contribution to the dielectric function is the so-called *orientational polarization* shown by the dashed line. It describes the contribution of freely rotatable permanent electric dipoles to $\varepsilon(\omega)$, i.e., a para-electric contribution. Freely rotatable

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0_25

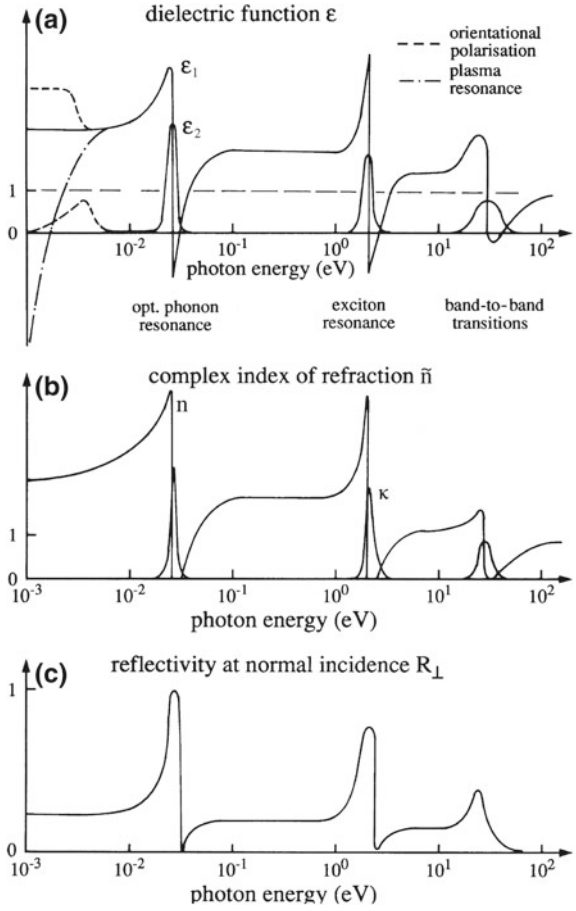


Fig. 25.1 Schematic overview from the IR to the UV of the spectra of the real and imaginary parts of $\varepsilon(\omega)$ (**a**), of $\tilde{n}(\omega)$ (**b**), and of the reflectivity (**c**) for a model semiconductor

means that the dipoles do not have a restoring force, but that an orientation of the dipoles which has been established, e.g., by an external field, decays after switching off of this field with a relaxation time τ by (thermal) collisions with quasi-particles in their environment. The contribution of the orientational polarization to ε_1 is constant for low frequencies as long as the dipoles can follow the driving field and it decays for $\omega\tau \geqslant 1$ as is usual for relaxation phenomena. The imaginary part has its maximum at $\omega\tau = 1$. This is reasonable if we assume that the friction losses of the rotating dipoles are proportional to the amplitude and the speed ω of the rotation. The orientational polarization is a very important effect for polar liquids (e.g., water) and organic solids containing some rotatable radicals. In semiconductors this effect is less common but is mentioned here for sake of completeness.

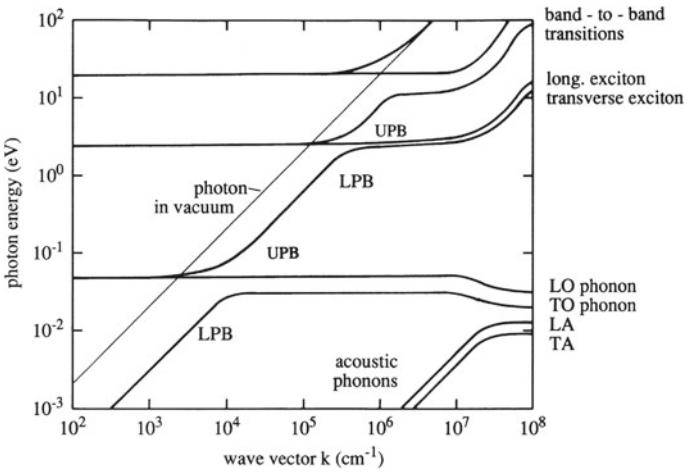


Fig. 25.2 Schematic overview of the real part of the dispersion relation of light in a semiconductor from the IR to the UV, neglecting damping

A schematic summary of *light propagation* in the model semiconductor is given by the dispersion relation depicted in Fig. 25.2,. Here we again include one phononic and two electronic resonances. For simplicity we consider zero damping and give only the real part of $\boldsymbol{k}$, but properly include spatial dispersion, i.e., the $\boldsymbol{k}$-dependence of the eigenfrequencies. The global *polariton dispersion* of Fig. 25.2 shows clearly the various resonances and how, above the highest eigenfrequency, the dispersion asymptotically approaches that of photons in vacuum.

Finally, the optical properties of a direct-gap semiconductor close to the bandgap are summarized in Fig. 25.3. Here we show schematically the *spectra* of the *absorption*, the *reflection*, and the *luminescence* at low temperature including free and bound excitons and donor–acceptor pairs, all with their LO-phonon replica. The exciton dispersion is plotted for comparison.

We end this summary and the first volume of this textbook by referring to a small collection of books covering aspects of semiconductor optics for further reading [69O1, 72O1, 72W1, 75B1, 91L1, 93O1, 93P1, 02D1, 02S1, 03B1, 04K1, 09H1, 09K1, 10Y1, 16G1, 17H1].

Problem

25.1 Try to identify the various schematic drawing in Figs. 25.1, 25.2 and 25.3 with real data in the relevant chapters.

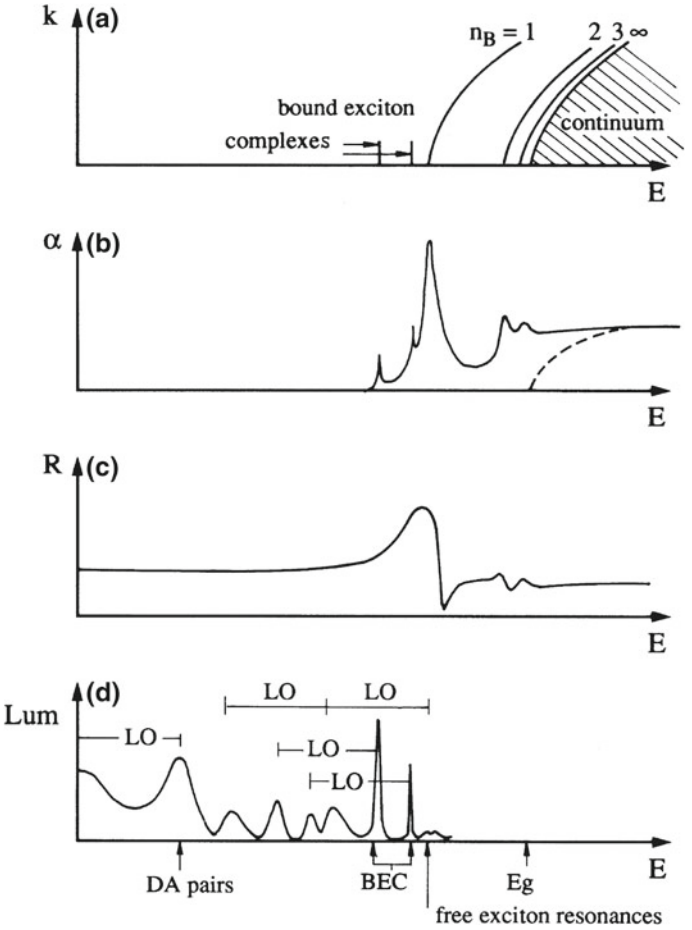


Fig. 25.3 Schematic drawing of the dispersion (**a**), and the spectra of the absorption (**b**), reflection (**c**) and luminescence (**d**) of a high-quality, direct-gap semiconductor at low temperatures in the region of the exciton resonances

References

[69O1] S. Nudelman, S.S. Mitra (eds.), *Optical Properties of Solids* (Plenum, New York, 1969)
[72O1] F. Abeles (ed.), *Optical Properties of Solids* (North-Holland, Amsterdam, 1972)
[72W1] F. Wooten, *Optical Properties of Solids* (Academic, New York, 1972)
[75B1] F. Bassani, G.P. Parravicini, *Electronic States and Optical Transitions in Solids* (Pergamon, Oxford, 1975)
[91L1] P.T. Landsberg, *Recombination of Semiconductors* (Cambridge University Press, Cambridge, 1991)
[93O1] F. Henneberger, S. Schmitt-Rick, E.O. Göbel (eds.), *Optics of Semiconductor Nanostructures* (Academic, Berlin, 1993)

[93P1] N. Peyghambarian, S.W. Koch, A. Mysyrowicz, *Introduction to Semiconductor Optics* (Prentice Hall, Engelwood Cliffs, 1993)
[02D1] M. Dressel, G. Grüner, *Eletrodynamics of Solids* (Cambridge University Press, Cambridge, 2002)
[02S1] W. Schäfer, M. Wegener, *Semiconductor Optics and Transport Phenomena* (Springer, Berlin, 2002)
[03B1] P.K. Basu, *Theory of Optical Processes in Semiconductors* (Oxford University Press, Oxford, 2013)
[04K1] H. Kalt, M. Hetterich (eds.), *Optics of Semiconductors and Their Nanostructures*. Springer Series in Solid-state Science, vol. 146 (Springer, Berlin, 2004)
[09H1] H. Haug, S.W. Koch, *Quantum Theory of the Optical and Electrical Properties of Semiconductors*, 5th edn. (World Scientific, Singapore, 2009)
[09K1] H. Kuzmany, *Solid-State Spectroskopie* (Springer, Berlin, 2009)
[10Y1] P.Y. Yu, M. Cardona, *Fundamentals of Semiconductors* (Springer, Berlin, 2010)
[16G1] M. Grundmann, *The Physics of Semiconductors* (Springer International Publishing, Berlin, 2016)
[17H1] C. Hamaguchi, *Basic Semiconductor Physics* (Springer International Publishing, Berlin, 2017)

Appendix A
Kramers–Kronig Relations

Abstract In this appendix we want to summarize in short some general relations between the real and imaginary parts of $\tilde{n}$ or ε derived from the theory of response functions.

The susceptibility χ and the dielectric function $\varepsilon = \chi + 1$ are *response functions* of the medium [72W1]. They describe the response (in this case the polarization density $\boldsymbol{P}$) to a stimulus (in this case the incident electric field $\boldsymbol{E}$) for the special case of an incident monochromatic wave with frequency ω. We now leave the description in the variables ω and $\boldsymbol{k}$ for a moment and consider their Fourier conjugate variables t and $\boldsymbol{r}$, i.e., time and real-space position vector.

The most general expression for a linear response function is

$$\frac{1}{\varepsilon_0}\boldsymbol{P}(\boldsymbol{r},t) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \chi\left(\boldsymbol{r},\boldsymbol{r}',t,t'\right)\boldsymbol{E}\left(\boldsymbol{r}',t'\right)\mathrm{d}t'\mathrm{d}\boldsymbol{r}' \,. \tag{A.1}$$

This equation describes the fact that the polarization $\boldsymbol{P}$ at point $\boldsymbol{r}$ and time t depends on the electric field at all other places and at all times.

Similar arguments can be given for the current density $\boldsymbol{j}\,(\boldsymbol{r},t)$ replacing the susceptibility by the electric conductivity $\sigma\left(\boldsymbol{r},\boldsymbol{r}',t,t'\right)$ (see (2.24) and (19.4)).

We are now going to simplify (A.1) in various steps. First we assume that the sample is homogeneous in time, i.e., its properties do not depend on t explicitly. Then χ depends only on the time difference $t - t'$. Since the medium consists of atoms, it is not homogeneous in space, but if we assume that the wavelengths considered for $\boldsymbol{E}\left(\boldsymbol{r}',t'\right)$ are much longer than the lattice constant, then in an analogous way χ depends only on $\boldsymbol{r} - \boldsymbol{r}'$. The latter is essentially the same assumption as has been made in the dipole approximation (6.23). Equation (A.1) then transforms into

$$\frac{1}{\varepsilon_0}\boldsymbol{P}(\boldsymbol{r},t) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \chi\left(\boldsymbol{r}-\boldsymbol{r}',t-t'\right)\boldsymbol{E}\left(\boldsymbol{r}',t'\right)\mathrm{d}t'\mathrm{d}\boldsymbol{r}' \,. \tag{A.2}$$

The response function χ is said to be *non-local*, i.e., the polarization density at $\boldsymbol{r}$ also depends on the electric field at other places $\boldsymbol{r}'$. In other words, a polarization

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0

density created at one place $\boldsymbol{r}'$ by the electric field at this place contributes at a later time to the polarization at another place $\boldsymbol{r}$. The finite propagation times of the fields obviously evoke *retardation* effects. This is just the phenomenon of *spatial dispersion*, as becomes clear by inspecting the models of Fig. 7.1 or of Sect. 8.4.

The integrals (A.1) and (A.2) are convolutions in space and in time. These convolutions simplify to products under Fourier transformation that are one dimensional in time and three dimensional in space [72W1]. Executing this Fourier transformation we obtain

$$\frac{1}{\varepsilon_0}\boldsymbol{P}\,(\boldsymbol{k},\omega) = \chi\,(\boldsymbol{k},\omega)\;\boldsymbol{E}\,(\boldsymbol{k},\omega)\;. \tag{A.3}$$

We now simplify the expression (A.3) in various steps. First we neglect spatial dispersion, i.e., we go from the situation in Fig. 8.8 back to Fig. 7.1. In this case the response function is local, i.e., $\boldsymbol{P}(\boldsymbol{r})$ depends only on $\boldsymbol{E}(\boldsymbol{r})$, reducing the dependence on $\boldsymbol{r}-\boldsymbol{r}'$ in (A.2) to a δ function $\delta\left(\boldsymbol{r}-\boldsymbol{r}'\right)$ and simplifying $\chi\,(\boldsymbol{k},\omega)$ to $\chi\,(\omega)$ leading to

$$\frac{1}{\varepsilon_0}\boldsymbol{P}(t) = \int\limits_{-\infty}^{+\infty}\chi(t-t')\boldsymbol{E}(t')\mathrm{d}t'\;. \tag{A.4}$$

Now we use a very important physical argument, namely *causality*. This argument is, in this context, valid both in classical physics and in quantum mechanics and means that the response $\boldsymbol{P}$ cannot come before the stimulus and thus

$$\chi(t-t') \equiv 0 \quad \text{for} \quad t' > t \tag{A.5}$$

or

$$\frac{1}{\varepsilon_0}\boldsymbol{P}(t) = \int\limits_{-\infty}^{t}\chi(t-t')\boldsymbol{E}(t')\mathrm{d}t'\;. \tag{A.6}$$

We now execute the above mentioned Fourier transformation with respect to time, resulting in

$$\boldsymbol{P}(\omega) = \int\limits_{-\infty}^{+\infty}\boldsymbol{P}(t)e^{\mathrm{i}\omega t}\mathrm{d}t\;, \tag{A.7a}$$

$$\boldsymbol{E}(\omega) = \int\limits_{-\infty}^{+\infty}\boldsymbol{E}(t)e^{\mathrm{i}\omega t}\mathrm{d}t\;, \tag{A.7b}$$

$$\chi(\omega) = \int\limits_{-\infty}^{+\infty}\chi(t-t')e^{\mathrm{i}\omega(t-t')}\mathrm{d}t\;. \tag{A.7c}$$

Here we see that a complex $\chi(\omega)$ results from a real $\chi(t-t')$. Inserting (A.6) into (A.7a) results in

$$\frac{1}{\varepsilon_0}\boldsymbol{P}(\omega) = \int_{-\infty}^{+\infty} e^{\mathrm{i}\omega t}\left[\int_{-\infty}^{t} \chi(t-t')\boldsymbol{E}(t')\mathrm{d}t'\right]\mathrm{d}t\ . \tag{A.8}$$

Introducing

$$1 = e^{-\mathrm{i}\omega t'}e^{\mathrm{i}\omega t'} \tag{A.9}$$

in the inner integral and rearranging the terms gives

$$\frac{1}{\varepsilon_0}\boldsymbol{P}(\omega) = \int \boldsymbol{E}(t')e^{\mathrm{i}\omega t'}\left[\int \chi(t-t')e^{\mathrm{i}\omega(t-t')}\mathrm{d}t\right]\mathrm{d}t' = \chi(\omega)\boldsymbol{E}(\omega)\ . \tag{A.10}$$

This is identical to (2.30).

With the knowledge of (A.10) we can apply Cauchy's theorem, which connects the real and imaginary parts of the Fourier transforms of analytic functions. This theorem leads to

$$\varepsilon_1(\omega) - 1 = \frac{1}{\pi}P\int_{-\infty}^{+\infty}\frac{\varepsilon_2(\omega')}{\omega'-\omega}\mathrm{d}\omega' \tag{A.11a}$$

and

$$\varepsilon_2(\omega) = -\frac{1}{\pi}P\int_{-\infty}^{+\infty}\frac{\varepsilon_1(\omega')-1}{\omega'-\omega}\mathrm{d}\omega'\,, \tag{A.11b}$$

where P in front of the integral means the principal value. Equation (A.11) can be rewritten as

$$\varepsilon_1(\omega) - 1 = \mathrm{Re}\{\chi(\omega)\} = \frac{2}{\pi}P\int_{0}^{\infty}\frac{\omega'\varepsilon_2(\omega)}{\omega'^2-\omega^2}\mathrm{d}\omega' \tag{A.12a}$$

and

$$\varepsilon_2(\omega) = \mathrm{Im}\{\chi(\omega)\} = -\frac{2\omega}{\pi}P\int_{0}^{\infty}\frac{\varepsilon_1(\omega')-1}{\omega'^2-\omega^2}\mathrm{d}\omega' \tag{A.12b}$$

since $\varepsilon(\omega) = \varepsilon^*(-\omega)$.

Similar relations hold for the phase $\phi(\omega)$ and amplitude $\rho(\omega)$ of the reflectivity $r(\omega)$ given for normal incidence by

$$r_\perp(\omega) = \frac{n(\omega)-1+\mathrm{i}\kappa(\omega)}{n(\omega)+1+\mathrm{i}\kappa(\omega)} = \rho(\omega)e^{\mathrm{i}\phi(\omega)}. \tag{A.13}$$

Compare (A.13) with the expression for $R_\perp$ in (5.19c). The relations for $\phi(\omega)$ and $\rho(\omega)$ read:

$$\ln \rho(\omega) = \frac{1}{\pi} P\int_{-\infty}^{+\infty} \frac{\phi(\omega')}{\omega' - \omega} \mathrm{d}\omega', \tag{A.14a}$$

$$\phi(\omega) = -\frac{1}{\pi} P\int_{-\infty}^{+\infty} \frac{\ln \rho(\omega')}{\omega' - \omega} \mathrm{d}\omega' = \frac{-2\omega}{\pi} P\int_{0}^{\infty} \frac{\ln \rho(\omega')}{\omega'^2 - \omega^2} \mathrm{d}\omega' . \tag{A.14b}$$

Relations between $n(\omega)$ and $\kappa(\omega)$, i.e., between the real and imaginary parts of $\tilde{n}(\omega)$ can be deduced from the relations

$$\mathrm{Im}\{\chi(\omega)\} = \varepsilon_2(\omega) = 2n(\omega)\kappa(\omega) \tag{A.15a}$$

$$\mathrm{Re}\{\chi(\omega)\} = \varepsilon_1(\omega) - 1 = n^2(\omega) - \kappa^2(\omega) - 1 . \tag{A.15b}$$

For systems, where $\chi(\omega)$ is small compared to unity, the following approximate relations hold

$$n(\omega) - 1 = \frac{2}{\pi} P\int_{0}^{\infty} \frac{\omega' \kappa(\omega')}{\omega'^2 - \omega^2} \mathrm{d}\omega' \tag{A.16a}$$

$$\kappa(\omega) = -\frac{2\omega}{\pi} P\int_{0}^{\infty} \frac{n_r(\omega')}{\omega'^2 - \omega^2} \mathrm{d}\omega' . \tag{A.16b}$$

Furthermore, the following relations hold:

$$n(\omega) = \frac{1 - \rho^2(\omega)}{1 + \rho^2(\omega) - 2\rho(\omega)\cos\phi(\omega)} , \tag{A.17a}$$

$$\kappa(\omega) = \frac{2 - \rho(\omega)\sin\phi(\omega)}{1 + \rho^2(\omega) - 2\rho(\omega)\cos\phi(\omega)} \quad \text{and} \tag{A.17b}$$

$$f = \frac{2}{\pi} \int_{0}^{\infty} \omega \varepsilon_2(\omega) \mathrm{d}\omega . \tag{A.17c}$$

The transverse and longitudinal eigenfrequencies can be deduced from the maxima of $\varepsilon_2(\omega)$ and of the so-called loss function $\mathrm{Im}\left\{-\frac{1}{\varepsilon(\omega)}\right\}$, respectively (see also [88R1]).

The relations (A.11), (A.12), (A.14), and (A.15) are known as *Kramers–Kronig relations*. They are of very general nature and rely only on causality and locality of the response. The two most important consequences are first, that if $\varepsilon_1(\omega)$ or $n(\omega)$ deviate in some frequency range from 1, then there must necessarily be absorption

structures somewhere, i.e. spectral regions must exist with $\varepsilon_2(\omega) \neq 0$ or $\kappa(\omega) \neq 0$ and vice versa. Secondly, if either the real or imaginary part of $\varepsilon(\omega)$, $\tilde{n}(\omega)$ or $r(\omega)$ is known over the whole spectral range, then the other variable can be calculated. Due to the denominator it is in practice sufficient to know the real or imaginary part over only a finite but not too small region around ω to be able to calculate the other part. The latter fact is used in experimental methods like ellipsometry (described in 16.2 and in [92H1]).

If spatial dispersion i.e. a non-local response is included, the Kramers–Kronig relations become more complicated. This topic is beyond the scope of this book and for further details we refer the reader to [70S1, 98D1].

For more details on the derivation and application of Kramers–Kronig Relations see e.g., [72W1, 92H1, 07M1, 07S1, 09K1, 10Y1].

Problems

A.1 If you are interested in the analysis of complex functions, derive the Kramers-Kronig relations from the properties of analytic complex functions $f(z)$ with $z \in \mathbb{C}$ and give the restrictions imposed on $f(z)$.

References

[70S1] T. Skettrup, Phys. Status Solidi **42**, 813 (1970)
[72W1] F. Wooten, *Optical Properties of Solids* (Academic, New York, 1972)
[88R1] H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*. Springer Tracts in Modern Physics, vol. 111 (Springer, Berlin, 1988)
[92H1] D.C. Hutchings et al., Opt. Quantum Electron. **24**, 1 (1992)
[98D1] B. Di Bartolo, in *Ultrafast Dynamics of Quantum Systems* (1997). NATO ASI Series B, vol. 372 (Plenum Press, New York, 1998)
[07M1] D. Meschede, *Optics, Light and Lasers*, 2nd edn. (Wiley, Weinheim, 2007)
[07S1] B.E.A. Saleh, M.C. Teich, *Fundamentals of Photonics*, 2nd edn. (Wiley, Hoboken, 2007)
[09K1] H. Kuzmany, *Solid-State Spectroscopy*, 2nd edn. (Springer, Berlin, 2009)
[10Y1] P.Y. Yu, M. Cardona, *Fundamentals of Semiconductors*, 4th edn. (Springer, Berlin, 2010)

Appendix B
Group Theory in Semiconductor Optics

By K. Hümmer, C. Klingshirn, and H. Kalt

Abstract In this chapter we will give an outline of group theory, its connection to quantum mechanics, and its particular applications in semiconductor optics. We will present only the most important rules and will in general give no proofs of the various relations. For more comprehensive treatments of group theory and more examples of applications we want to refer to text books on group theory and on crystallography, e.g., [60H1, 64K1, 64T1, 67S1, 74B1, 88L1, 90I1, 92I1, 94H1, 10Y1, 01L1, 18H1]. Very good collections of tables are found in the books by Koster et al. [63K1] and by Ludwig et al. [88L1], from which most of the tables in this Appendix are taken. In the following sections we will describe the basic properties of groups in Sect. B.2 and then introduce the terms *representations* and *characters* (Sect. B.3). Then we present the central part of this chapter, namely the implications of symmetry, i.e., group theory on the Hamiltonian in Sect. B.4. Finally we give applications both of general nature and others which are more or less specific to semiconductors in Sects. B.5 and B.6. The aim of this chapter is not to give the reader a deep insight into group theory and its implications. The idea is to gain a qualitative understanding of symmetry considerations and how they affect the optical properties of solid. Also the reader should understand the meaning of group theory notations like the Γ_i that are met in plots of the semiconductor band structure as in Chap. 15. We also want to enable him or her to calculate simple *selection rules* for optical transitions.

B.1 Noether's Theorem Applied to Crystalline Solids

Noether's theorem has been introduced in Sect. 5.1.3 and has been referred to several times in this book. It states that a conservation law follows from every invariance of the Hamiltonian. The classic examples are the following:

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0

- If H is invariant under an infinitesimally shift in time, then the total energy of the system described by this Hamiltonian is conserved.
- If H does not depend on a particular spatial coordinate x_i, i.e., if it is invariant under an infinitesimal translation along x_i, then the component of the momentum p_i in the direction x_i is conserved.
- If H has spherical symmetry (as for a single atom, for example) it is invariant with respect to an infinitesimal rotation $\mathbf{d}\phi$ around any axis, and a conservation law for the angular momentum follows for L^2 and for one component of $\boldsymbol{L}$, e.g., L_z.

In a crystal the first condition remains valid in the absence of explicitly time-dependent (perturbative) terms, i.e., energy is still conserved.

The second condition is no longer valid, but is replaced by invariance with respect to translations by integer multiples of the $\boldsymbol{a}_i$ (Sect. 9.2). We have already presented this subject and its consequences in some detail, e.g., in connection with light propagation through interfaces in Sect. 5.1.3, with the reciprocal lattice in Chap. 9, with the propagation of phonons in Chaps. 10 and 11, or Bloch's Theorem in Sect. 14.2. We will return to it briefly in this chapter, too.

The third condition concerning infinitesimal rotations likewise no longer applies to crystals. Instead we have, depending on the crystal structure, only invariance with respect to rotations around two, three, four, and sixfold symmetry axes. This means that strictly speaking $\boldsymbol{L}$ is not a good quantum number in crystals, but later, in connection with compatibility relations, we shall see that arguments based on $\boldsymbol{L}^2$ and/or L_z can still sometimes be used, but with considerable care.

B.2 Basics of Abstract Group Theory for Crystalline Solids

B.2.1 Definition of a Group

A *group* G is a set of *elements* $\{X_i\}$ and a connection or composition between the elements which is often called *multiplication* with the following properties:

- *Closure*: if X_i and X_j are elements of G $(X_i, X_j \in G)$ then the *product* is element of G

$$X_i X_j = X_k \in G\,. \tag{B.1a}$$

- *Associative law*:

$$X_i \left(X_j X_k\right) = \left(X_i X_j\right) X_k\,. \tag{B.1b}$$

- There is a *neutral element*, also called *identity operation* or *identity mapping* E in G with

$$E X_i = X_i E = X_i \quad \text{for} \quad X_i \in G\,. \tag{B.1c}$$

- There is an *inverse element* X_i^{-1} for each $X_i \in G$ with

$$X_i^{-1} X_i = X_i X_i^{-1} = E\,. \tag{B.1d}$$

The number g of different elements in a group gives the *order* of the group. There are finite and infinite groups. If the *commutative law* holds, i.e., if

$$X_i X_j = X_j X_i \quad \text{for all} \quad X_i X_j \in G \tag{B.2}$$

the group is said to be *Abelian*.

A group is called *cyclic* if it can be generated by one of its elements $G = \{X, XX = X^2, XXX = X^3, \ldots\}$. If one of these elements is the unit element, $X^k = E$, then G is finite and cyclic of order k.

A subset H of a group G is called a *subgroup* of G if H obeys the group axioms. The whole group G and E are called trivial or improper subgroups; the others are called proper subgroups and one writes $H < G$.

A set of elements from which the complete group can be obtained by composition, is called a set of generators of G. A cyclic group obviously has exactly one such generator.

B.2.2 Classification of the Group Elements

A distribution of the elements into subsets such that each element belongs to exactly one subset, i.e., two different subsets have no element in common, is called a *classification*. There are two important methods to classify the elements of a group: (1) *coset* decomposition and (2) partition into *conjugacy classes*.

1. A group can be decomposed into left or right cosets relative to a subgroup, say $H = \{Y\}$, by multiplying each $X_i \in G$ with the elements of H from left $X_i\{Y\}$ or right $\{Y\}X_i$, respectively. Exactly one of these cosets is a group, namely the subgroup H.
2. An element X_j is called *conjugate* to X_i, if an element $X_k \in G$ exists for which the following holds:

$$X_k^{-1} X_i X_k = X_j\,. \tag{B.3}$$

   If X_j is conjugated to X_i then X_i is also conjugated to X_j with the element $X_k^{-1} \in G$, which must exist. The subset of the elements of G, which are conjugate to X_i, X_k running through G, is the conjugacy class of X_i. The number of elements in a conjugacy class is called its length l, which may be different for different conjugacy classes, but is always a factor of g. Elements of the same conjugacy class have the same order with respect to their cyclic group.

An element for which

$$X_k^{-1} X_i X_k = X_i \tag{B.4}$$

holds for all $X_k \in G$ is called *self-conjugate* in G. Therefore, in Abelian groups each $X_i \in G$ forms a conjugacy class by itself of length $l = 1$.

Conjugacy is also defined for subgroups. Of special interest are self-conjugate subgroups, also called normal or invariant subgroups for which holds: $X_k^{-1}\{Y\}X_k = \{Y\}$ for all $X_k \in G$ where $\{Y\} = H$ is a proper subgroup of $G : H < G$. Then, $\{Y\}X_k = X_k\{Y\}$, i.e. for a normal subgroup the right and left coset decomposition coincide.

The cosets of G relative to a normal subgroup H are elements of a group, called the *factor group* F, sometimes denoted as $F = G/H$. The unit element of F is the normal subgroup H.

B.2.3 Isomorphism and Homomorphism of Groups

Isomorphism is a special case of the more general concept of *homomorphism*. An isomorphism is a one-to-one correspondence between the elements of groups, which is preserved during multiplication. Two groups G and G' with elements X_i and X_i' are isomorphic if (after suitable ordering of their elements) there is a reversible mapping of G onto G': $X_i \leftrightarrow X_i'$, and for each pair $X_i, X_j \in G$ the product of the images X_i' $X_j' \in G'$ is equal to the image of the product: $X_i' X_j' = (X_i X_j)'$.

The isomorphism yields a classification of all possible groups into isomorphism classes, which is sometimes called an *abstract group*. The abstract group displays the common group theoretical features of isomorphic groups independent of the realization of their elements.

In contrast, a homomorphism maps G onto a possibly smaller group G' and is then an irreversible, many-to-one correspondence between G and G', which is preserved during multiplication. More formally this reads: $X_i' X_j' = (X_i X_j)'$.

B.2.4 Some Examples of Groups

We now want to give some examples for groups starting from simple examples for numbers, proceeding to symmetry operations of simple physical and geometric objects, and finally to space groups of crystals.

Example (1)
The positive and negative integers including zero $\{0, \pm 1, \pm 2, ...\}$ with normal addition as connection. They form an infinite Abelian group.

Example (2)
The positive rational numbers p/q with $p = 0, 1, 2, 3, \ldots$ and $q = 1, 2, 3, \ldots$ with normal multiplication as connection. They also form an infinite Abelian group.

Example (3)
Symmetry operations of any physical object, e.g., a atom, a molecule or a crystalline solid, by definition map an object onto itself. Symmetry operations are thus geometric mappings that leave all distances invariant, so-called isometries or rigid motions. They form a group, called the *symmetry group* of the object. In any symmetry group the unit element is the identity mapping that leaves every point fixed.

Example (4)
In the following example we use a simple geometric model (an equilateral triangle) in order to illustrate the definitions of abstract group theory given above.

The symmetry operations that map an equilateral triangle onto itself form a group (Fig. B.1). The symmetry operations are: E, the unit element, which simply leaves the triangle unchanged; J and K are rotations of $+120°$ and $240° = -120°$, respectively, around the axis through the center of the triangle and normal to it; L, M, and N are rotations of 180° around the axes a–c. In the latter case the upper and lower faces of the triangle must be indistinguishable. Alternatively, there can be mirror planes normal to the triangle that contain the a-, b- or c-axis.

The symmetry elements at which the rotations are carried out are called either threefold or twofold axes as the cyclic group of J is of order 3: $JJJ = J^3 = E$, $J^2 = J^{-1} = K$. That of L is of order 2: $LL = L^2 = E, L = L^{-1}$. In crystallographic group notations they are denoted by 3 and 2, a mirror plane by m. In 3-dimensional

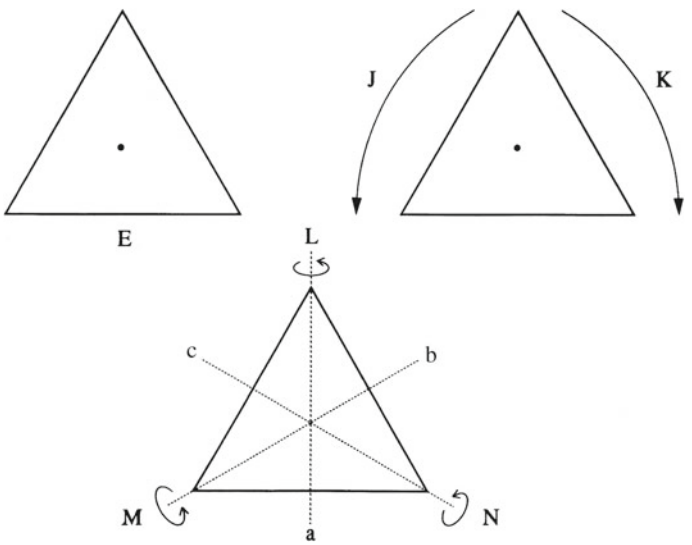


Fig. B.1 The symmetry operations which map an equilateral triangle into itself

Table B.1 Multiplication table of group D_3. We give the names of the elements and the operations corresponding to Fig. B.1

	E	$J(+120)$	$K(-120)$	$L(180_a)$	$M(180_b)$	$N(180_c)$
E	E	$J(+120)$	$K(-120)$	$L(180_a)$	$M(180_b)$	$N(180_c)$
$J(+120)$	$J(+120)$	$K(-120)$	E	$N(180_c)$	$L(180_a)$	$M(180_b)$
$K(-120)$	$K(-120)$	E	$J(+120)$	$M(180_b)$	$N(180_c)$	$L(180_a)$
$L(180_a)$	$L(180_a)$	$M(180_b)$	$N(180_c)$	E	$J(+120)$	$K(-120)$
$M(180_b)$	$M(180_b)$	$N(180_c)$	$L(180_a)$	$K(-120)$	E	$J(+120)$
$N(180_c)$	$N(180_c)$	$L(180_a)$	$M(180_b)$	$J(+120)$	$K(-120)$	E

space these are the symmetry elements of a trigonal trapezohedron, denoted as $D_3 = 32$ (Schönflies and Hermann–Mauguin notation) and of a trigonal pyramid, denoted as $C_{3v} = 3m$.

The composition (multiplication) of the symmetry operations is to successively perform these operations in a well-defined order. A convenient way to display a finite group (of small order) is the *multiplication* or *group table*. For our current example it is displayed in Table B.1). It is a square array, where the product $X_i X_j$ is listed at the intersection of the row of operation X_i and the column of operation X_j. The notation $X_i X_j$ means to first apply X_j and then X_i. In every row and in every column of the group table, every operation of the group appears once and only once.

The group table of D_3 and C_{3v}, as well as the group of all permutations of three different symbols, say (abc), can be made identical apart from the names or symbols of the operations: $E = (abc)$, $J = (cab)$, $K = (bca)$, $L = (acb)$, $M = (cba)$, $N = (bac)$. There exists a one-to-one correspondence between the operations of the three groups, i.e., they are isomorphic. They are of order $g = 6$.

The fact that the group table is not symmetric with respect to the main diagonal shows that the group is non-Abelian. In fact, it is known as the non-Abelian group of lowest order.

The following proper subgroups can be deduced as can be seen from an inspection of the group table: $\{E, J, K\}$ with $g = 3$ and $\{E, L\}, \{E, M\}, \{E, N\}$ with $g = 2$.

The coset decomposition relative to the subgroup $\{E, J, K\}$ gives one more coset $\{L, M, N\}$, i.e., left and right cosets are identical, and $\{E, J, K\}$ is a normal or invariant subgroup. The coset decomposition relative to the subgroup $\{E, L\}$ gives two additional left cosets $\{J, N\}, \{K, M\}$ and two more right cosets $\{J, M\}, \{K, N\}$ as can be seen from an inspection of the group table.

In the abstract group mentioned above there are three conjugacy classes, or in short simply C_i classes, as can be verified using the group table: $C_1 = \{E\}$ with $l = 1$; $C_2 = \{J, K\}$ with $l = 2$; $C_3 = \{L, M, N\}$ with $l = 3$. For example, the operations J and K belong to the same class since

$$N^{-1} J N = NM = K \quad \text{with} \quad N^{-1} = N\,. \tag{B.5}$$

Geometrically, conjugacy means the search for symmetrically equivalent symmetry elements. They must be of the same type, or from the point of view of group theory, their cyclic groups must be of the same order. In fact, the two rotations J and K around the threefold axis are symmetrically equivalent through the three rotations L, M, N around the twofold axes or through the respective mirror operations and vice versa. The rotations around the twofold axes (or the mirror operations) are symmetrically equivalent through the rotations around the threefold axes.

Example (5)

The set of all symmetry operations (isometries) of an idealized infinite crystal structure forms a crystallographic *space group*. In three-dimensional space we distinguish 230 types of crystallographic space groups of order $g = \infty$.

The most characteristic feature of a crystal is that its structure can be embedded into a lattice. The existence of a lattice, i.e. the presence of three-dimensional periodicity, gives rise to diffraction effects, as first shown by Laue, Friedrich and Knipping. In terms of group theory this corresponds to the existence of a discrete translation subgroup T of every space group G.

The symmetry operations of T are translations by lattice vectors $\boldsymbol{R}_n$. Choosing a suitable basis of lattice vectors, a so-called primitive basis, each lattice vector is represented by a linear combination of the basis vectors with integer coefficients (see also Sect. 9.1):

$$\boldsymbol{R}_n = \sum_i n_i \boldsymbol{a}_i \,. \tag{B.6}$$

Isometries are actually special affine mappings described by a system of equations

$$\widetilde{\boldsymbol{x}} = \boldsymbol{D}\boldsymbol{x} + \boldsymbol{R} = (\boldsymbol{D}, \boldsymbol{R})\boldsymbol{x} \,. \tag{B.7}$$

The pair $(\boldsymbol{D}, \boldsymbol{R})$ consists of the (3×3)-matrix $\boldsymbol{D}$, the rotational part, and the (3×1)-column $\boldsymbol{R}$, the translational part. For isometries, i.e., distance, preserving mappings, the condition $det\ \boldsymbol{D} = \pm 1$ must be obeyed.

The lattice of translation vectors of a crystal imposes restrictions on the isometries occurring in space groups. The rotational matrix $\boldsymbol{D}$ has to obey $\boldsymbol{D}^k = \boldsymbol{E}$, with the order $k = 2, 3, 4, 6$. That means a crystal only allows rotations by multiples of 180, 120, 90, and 60°.

Thus, in crystallographic space groups the following symmetry operations occur. Their symbols are given in the Hermann–Mauguin notation:

(a) *Identity* $(\boldsymbol{E}, \boldsymbol{0})$. Symbol 1: $\boldsymbol{E}$ is the unit matrix.
(b) *Translations* by lattice vectors $(\boldsymbol{E}, \boldsymbol{R}_n)$.
(c) *Rotations* and *screw rotations* $det\ \boldsymbol{D} = +1$

(i) Rotations. Symbols $k = 2, 3, 4, 6$ according to their order k; for each one exists a line of (invariant) fixed points, the rotation axis.
(ii) Screw rotations. Symbols k_n with $n < k$: $(\boldsymbol{D}, \boldsymbol{R})^k = (\boldsymbol{E}, \boldsymbol{R}_n)$ must hold where $\boldsymbol{R}_n$ is a lattice translation; $(\boldsymbol{D}, \boldsymbol{R})$ can be decomposed into a rotation

around an axis (screw axis) and a translation by the screw vector parallel to this axis $\boldsymbol{R}_{\parallel} = \frac{n}{k}\boldsymbol{R}_0$, where $\boldsymbol{R}_0$ is the shortest lattice vector along the screw axis.

(d) *Inversion* $\boldsymbol{D} = -\boldsymbol{E}$. Symbol $\bar{1}$: exactly one point is fixed, the center of inversion or center of symmetry.
(e) *Rotoinversions* $det\, \boldsymbol{D} = -1$. Symbols $\bar{k}, k \neq 2$: it may be decomposed into a rotation followed by an inversion.
(f) *Reflection* and *glide reflection* $det\, \boldsymbol{D} = -1$, $\boldsymbol{D}^2 = \boldsymbol{E}$:

(i) Reflection $(\boldsymbol{D}, \boldsymbol{R})^2 = (\boldsymbol{E}, \boldsymbol{0})$. Symbol m: the fixed points form a plane at which reflection of space is performed, the so-called mirror plane. It is a special rotoinversion $\bar{2} = m$ where the axis stands normal to the mirror plane.
(ii) Glide reflection $(\boldsymbol{D}, \boldsymbol{R})^2 = (\boldsymbol{E}, \boldsymbol{R}_n)$. Symbols a, b, c, n, d: may be decomposed into a reflection through a plane (glide plane) and a translation by the glide vector parallel to this plane.

As already mentioned above, the subset of isometries $\{(\boldsymbol{E}, \boldsymbol{R}_n)\}$, i.e., translations by lattice vectors, forms a subgroup T of the space group G. It is normal or invariant since the left and right coset decomposition coincide. The symmetry operations in each coset relative to T have the rotational part $\boldsymbol{D}$ in common since the product of any space group operation $(\boldsymbol{D}, \boldsymbol{R})$ and any lattice translations $(\boldsymbol{E}, \boldsymbol{R}_n)$ does not change the rotational part $\boldsymbol{D}$. The factor group $F = G/T$ where the cosets are the elements of F, is isomorphic to the point group P of the space group G. This point group P of the crystal can thus be obtained directly from the coset decomposition relative to T by considering only the rotational parts given by (3×3)-matrices.
It is important to know the *point group* P or the crystal class of the space group G of the crystal structure. It describes the symmetry of the form of the crystal, the symmetry of the lattice and thus of the reciprocal lattice. The latter determines the symmetry of a diffraction pattern.
There is one further important aspect of this coset decomposition, namely the denotation of space groups by means of the *Hermann–Mauguin symbols* used in crystallography. The first character is the letter of the type of the Bravais lattice of the structure given by the translation subgroup T. Here P means primitive; A, B and C mean one face is centered; I means body centered; F means all faces are centered; R means special centering type in trigonal crystal classes. Then follows a finite set of symbols (maximum three) of generators that when combined with $\{T\}$ permits the generation of all space group operations.

In the following we give some examples of space groups for important types of semiconductor structures.

(a) *Wurtzite structure*
Prototype: ZnO; other examples CdS, CdSe, GaN.

$$spacegroup G\colon P6_3mc; \quad pointgroup P\colon 6mm; \quad \text{Schönflies notation: } C_{6v}\,.$$

The point group symbol P in the Hermann–Mauguin notation is simply derived by replacing the screw axes by their rotation axes and the glide planes by their mirror planes.
The point group $6mm$ belongs to the hexagonal crystal system. It is of order $g = 12$ with the following symmetry operations. Rotation axis 6 contains the operations $6^+ =$ rotation by 60°, $6^- =$ rotation by $-60°$, $3^+ =$ rotation by 120°, $3^- =$ rotation by $-120°$, and $2 =$ rotation by 180°. There are two triples of mirror planes. In each triple the symmetrically equivalent mirror planes enclose an angle of 60°. However, the mirror planes of different triples are rotated by 30° against each other. Thus, they are not symmetrically equivalent (there is no symmetry operation rotation by 30°) and they belong to different conjugacy classes, which is why the Hermann–Mauguin symbol $6mm$ has two sets of non-equivalent mirror planes indicated. Thus, in the single group there are six (conjugacy) classes denoted in the Schönflies notation as (see Table B.7): E, $2C_6$, $2C_3$, C_2, $3\sigma_d$, $3\sigma_v$.

(b) *Zincblende structure*
Prototype: ZnS, other examples are III–V-compounds like GaAs, InP, or I–VII compounds like the Cu-halides or AgBr.

$$spacegroup G\colon F\bar{4}3m; \quad pointgroup P\colon \bar{4}3m; \quad \text{Schönflies notation: } T_d\,.$$

The point group $\bar{4}3m$ belongs to the cubic crystal system. It is of the order $g = 24$. It is non-centrosymmetric. The classes of the single group contain (see Table B.10), apart from the identity class, the operations 3^+ and 3^- of four diagonal threefold axes ($8C_3$) of a cube, twofold operations rotation by 180° of three axes $2 = \bar{4}^2(3C_2)$, the operations $\bar{4}^1 = \bar{4}^+$ and $\bar{4}^3 = \bar{4}^-(6S_4)$ of the three $\bar{4}$ rotoinversion axes along the axes of a cube and twofold operations at six diagonal mirror planes ($6\sigma_d$).

(c) *Diamond structure*
Prototype: C, other examples Ge, Si, Cu_2O, NaCl.

$$spacegroup G\colon Fd\bar{3}m; \quad pointgroup P\colon m\bar{3}m; \quad \text{Schönflies notation: } O_h\,.$$

The point group $m\bar{3}m$ is centro-symmetric and belongs to the cubic crystal system. It is of order $g = 48$. It can be considered as the direct product of $\bar{4}3m = T_d$ and the point group $\bar{1} = C_i$. Thus, we also have the classes (see Table B.15) C_i, $6C_4$ (since $\bar{4} \times \bar{1} = 4$), $8C_{3i}$ (since $3 \times \bar{1} = \bar{3} = C_{3i}$) and $3\sigma_h$ (since $2 \times \bar{1} = m$).

The *outer* or *direct product* of two groups $G = \{X_i\}$ and $G' = \{X'_i\}$ is again a group G'' which consists of all ordered pairs X_i, X'_j. The definition of the connection of the elements of G'' is

$$(X_i, X'_j)(X_k, X'_l) = (X_i X_k, X'_j X'_l), \tag{B.8}$$

and the order of G'' is $g \cdot g'$.

B.3 Theory of Representations and of Characters

Any set of elements which fulfills the multiplication table of a group is called a *representation* of this group. The operations in Fig. B.1 can be considered as a representation of D_3.

In a stricter sense, to which we shall stick from now on, we consider a representation to be a set of $(n \times n)$ matrices $(n = 1, 2, 3, \ldots)$ which fulfills the multiplication table with normal matrix multiplication as the connection.

The name of a certain representation is given by a label α used as an index. With $\mathbf{\Gamma}_\alpha$ we denote a set of g matrices that fulfill the multiplication table of group G. $\mathbf{\Gamma}_\alpha(X_i)$ is the matrix that represents the element X_i of G and $\mathbf{\Gamma}_\alpha(X_i)_{jk}$ is the element of this matrix in row j and column k.

The normal matrix multiplication then results in

$$\mathbf{\Gamma}(X_i, X_j)_{kl} = \sum_h \mathbf{\Gamma}_\alpha(X_i)_{k,h} \mathbf{\Gamma}_\alpha(X_j)_{h,l}. \tag{B.9}$$

We introduce n_α to denote the *dimension* of the $(n_\alpha \times n_\alpha)$ matrices of a certain representation $\mathbf{\Gamma}_\alpha$; n_α is the same for all matrices of this particular representation.

There is an infinite number of representations for one group. If $\mathbf{\Gamma}_\alpha$ is a representation and $\boldsymbol{X}$ is a non-singular matrix of the same dimension n_α then $\{\boldsymbol{X}^{-1}\mathbf{\Gamma}_\alpha\boldsymbol{X}\}$ is again a representation. In the following we will distill from this infinite number of representations some with special properties (the so-called *irreducible representations*) and present some of their characteristic properties.

Table B.2 gives four examples of representations of group D_3, two one-dimensional ones, a two- and a three-dimensional one. A representation $\mathbf{\Gamma}_\alpha$ is called *reducible* if a non-singular matrix $\boldsymbol{Y}$ can be found such that all matrices $\boldsymbol{Y}^{-1}\mathbf{\Gamma}_\alpha(\boldsymbol{X}_i)\boldsymbol{Y}$ acquire the same block form. This means there are square sub-matrices and otherwise zeros; see Fig. B.2. From the properties of matrix multiplication it follows immediately that only submatrices at identical positions are connected with each other in multiplication. Consequently the submatrices of all X_i which are located in the same position, alone form a representation.

This means, that a representation $\mathbf{\Gamma}_\alpha$ which can be brought into such a block form of submatrices is called reducible. The representation $\mathbf{\Gamma}_4$ in Table B.2 is evidently reducible. It is already expressed in the desired block form. Representation $\mathbf{\Gamma}_\alpha$, for

Table B.2 Four matrix representations of the group D_3

	E	J	K	L	M	N
Γ_1	1	1	1	1	1	1
Γ_2	1	1	1	-1	-1	-1
Γ_3	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}$	$\frac{1}{2}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$
"Γ_4"	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

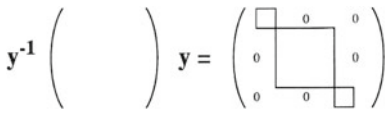


Fig. B.2 The reduction of a reducible representation into irreducible ones

which no matrix $\boldsymbol{Y}$ can be found to bring all $\boldsymbol{\Gamma}_\alpha(X_i)$ into the same block form are called irreducible. The representations $\boldsymbol{\Gamma}_1 - \boldsymbol{\Gamma}_3$ of Table B.2 are irreducible. We will focus in the following on irreducible representations.

Two irreducible representations $\boldsymbol{\Gamma}_\alpha$, and $\boldsymbol{\Gamma}_\beta$ are called *equivalent*, if there exists a non-singular matrix $\boldsymbol{S}$ with

$$\boldsymbol{S}^{-1}\boldsymbol{\Gamma}_\alpha(\boldsymbol{X}_i)\boldsymbol{S} = \boldsymbol{\Gamma}_\beta(\boldsymbol{X}_i) \quad \text{for all} \quad \boldsymbol{X}_i \in \boldsymbol{G}\,. \tag{B.10}$$

With this definition of equivalence we arrive at a very important statement: There exist exactly as many non-equivalent (i.e., different) irreducible representations of a group as the group has classes C_i.

For D_3 we found three classes (see above). Consequently D_3 has three non-equivalent irreducible representations. These are the $\boldsymbol{\Gamma}_1$, $\boldsymbol{\Gamma}_2$ and $\boldsymbol{\Gamma}_3$ in Table B.2. The irreducible representation in which all elements are represented by 1 is called the *trivial representation* and is always called $\boldsymbol{\Gamma}_1$. The representations of the identity element are always unit matrices with 1 on the main diagonal and zeros otherwise.

For irreducible representations $\boldsymbol{\Gamma}_\alpha$ and $\boldsymbol{\Gamma}_\beta$ there exists an orthogonality relation

$$\sum_{X_i} \boldsymbol{\Gamma}_\alpha(X_i)_{kp}\boldsymbol{\Gamma}_\beta(X_i^{-1})_{ql} = \frac{g}{n_\alpha}\delta_{\alpha\beta}\delta_{kl}\delta_{pq} \tag{B.11a}$$

with δ_{kl} and δ_{pq} being Kronecker's symbols and

$$\delta_{\alpha\beta} = \begin{cases} 0 & \text{if } \boldsymbol{\Gamma}_\alpha \text{ and } \boldsymbol{\Gamma}_\beta \text{ are not equivalent} \\ 1 & \text{if } \boldsymbol{\Gamma}_\alpha \text{ and } \boldsymbol{\Gamma}_\beta \text{ are identical} \\ \neq 0 & \text{but undefined if } \boldsymbol{\Gamma}_\alpha \text{ and } \boldsymbol{\Gamma}_\beta \text{ are equivalent} \end{cases} \tag{B.11b}$$

Table B.3 The character table of the group D_3

	Classes of elements		
Irreducible representations	E	$2C_3$	$3C_2$
$\mathbf{\Gamma}_1$	1	1	1
$\mathbf{\Gamma}_2$	1	1	–1
$\mathbf{\Gamma}_3$	2	–1	0

A characteristic quantity which is identical for all equivalent irreducible representations is the *trace* of the matrices

$$\chi_\alpha(X_i) = Tr\,\mathbf{\Gamma}_\alpha(X_i) = \sum_j \mathbf{\Gamma}_\alpha(X_i)_{jj}\,. \tag{B.12}$$

These traces are called *characters*. Two equivalent representations have the same characters since

$$\mathrm{Tr}\mathbf{\Gamma}_\alpha(X_i) = Tr\,\boldsymbol{S}^{-1}\mathbf{\Gamma}_\alpha(X_i)\boldsymbol{S}\,. \tag{B.13}$$

We can thus write down the characters of all non-equivalent irreducible representations in a table. Since the characters of conjugate elements are the same, as seen from a comparison of (B.3) and (B.13), it is sufficient to give the characters for the different classes. This results in a square scheme of numbers. Table B.3 gives as a first example the *character table* for the group D_3. It is identical, by the way, to that of the group C_{3v} mentioned above.

As already mentioned, $\mathbf{\Gamma}_1$, in the first line is always the trivial representation, thus the first column gives the dimensionality n_α of the representation $\mathbf{\Gamma}_\alpha$. The nomenclature $\chi_\alpha(C_i)$ means the character of the irreducible representation a for the class C_i.

Introducing the number h_i of elements of the class C_i, with a total number r of classes and of irreducible representations in the group, we find the following relations for the irreducible representations

$$\sum_{i=1}^{r} h_i\chi_\alpha(C_i)\chi^*_\beta(C_i) = g\delta_{\alpha\beta}\,, \tag{B.14a}$$

$$\sum_{\alpha=1}^{r} h_i\chi_\alpha(C_i)\chi^*_\alpha(C_j) = g\delta_{ij}\,. \tag{B.14b}$$

Two representations $\mathbf{\Gamma}_\alpha$ and $\mathbf{\Gamma}_\beta$ are equivalent if $\chi_\alpha(X_i) = \chi_\beta(X_i)$ for all X_i. If a representation $\mathbf{\Gamma}$ is given, it contains the irreducible representation $\mathbf{\Gamma}_\alpha$ p_α times in the sense of a decomposition into block form:

$$\mathbf{\Gamma} = \sum_\alpha p_\alpha\mathbf{\Gamma}_\alpha \quad \text{with} \quad p_\alpha = \sum_{X_i} \chi(X_i)\chi^*_\alpha(X_i)\,. \tag{B.15}$$

This means that it is not necessary in practice to try out all possible non-singular matrices since one can immediately deduce from (B.15) the decomposition of a given representation into irreducible ones. The reducible representation "$\mathbf{\Gamma}_4$" in Table B.2 can consequently be decomposed into

$$\mathbf{\Gamma}_4 = \mathbf{\Gamma}_2 \oplus \mathbf{\Gamma}_3 \,. \tag{B.16}$$

This process also explains the use of the symbol $\oplus$.

The *direct product* $\mathbf{\Gamma}_\alpha \otimes \mathbf{\Gamma}_\beta$ of two representations $\mathbf{\Gamma}_\alpha$ and $\mathbf{\Gamma}_\beta$ with dimensions n_α and n_β, respectively, is a new representation of dimension $n_\alpha \cdot n_\beta$. The new set of matrices is obtained in the following way

$$\mathbf{\Gamma}_\alpha(X_i) \otimes \mathbf{\Gamma}_\beta(X_i) = \begin{pmatrix} \mathbf{\Gamma}_\alpha(X_i)_{11}\mathbf{\Gamma}_\beta(X_i) & \dots & \mathbf{\Gamma}_\alpha(X_i)_{1n_\alpha}\mathbf{\Gamma}_\beta(X_i) \\ \vdots & & \\ \mathbf{\Gamma}_\alpha(X_i)_{n_\alpha 1}\mathbf{\Gamma}_\beta(X_i) & \dots & \mathbf{\Gamma}_\alpha(X_i)_{n_\alpha n_\alpha}\mathbf{\Gamma}_\beta(X_i) \end{pmatrix} \tag{B.17}$$

The direct product of two representations is again a representation of the group. The direct product of two irreducible representations can be reducible or irreducible.

For the characters of the direct product one finds

$$\chi\left(\mathbf{\Gamma}_\alpha \otimes \mathbf{\Gamma}_\beta\right) = \chi_\alpha \cdot \chi_\beta \,. \tag{B.18}$$

The decomposition of the products of the irreducible representations of a group is usually given in tables:

$$\mathbf{\Gamma}_\alpha \otimes \mathbf{\Gamma}_\beta = \sum_\gamma g_{\alpha\beta\gamma}\mathbf{\Gamma}_\gamma \tag{B.19}$$

with

$$g_{\alpha\beta\gamma} = \frac{1}{g}\sum_{X_i} \chi_\alpha(X_i)\chi_\beta(X_i)\chi^*_\gamma(X_i) \,. \tag{B.20}$$

We give the direct products of the irreducible representations of D_3 in Table B.4.

The criterion for a given representation to be irreducible is

$$\mathbf{\Gamma}_\alpha \text{ irred} \quad \Leftrightarrow \quad \sum_{X_i} |\chi_\alpha(X_i)|^2 = g \,. \tag{B.21}$$

Table B.4 Multiplication table for the irreducible representations of the group D_3

D_3	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_3$
$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_3$
$\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_3$
$\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_1 + \mathbf{\Gamma}_2 + \mathbf{\Gamma}_3$

Table B.5 Compatibility table of the irreducible representations of the group D_3 with the irreducible representations of its proper subgroups C_2 and C_3

D_3	Γ_1	Γ_2	Γ_3
C_3	Γ_1	Γ_1	$\Gamma_2 + \Gamma_3$
C_2	Γ_1	Γ_2	$\Gamma_1 + \Gamma_2$

A representation of a group is also a representation of each of its subgroups. An irreducible representation of a group can be a reducible or irreducible representation of the subgroup. Which possibility applies can either be checked with formula given above or it can be found in the tables. In Table B.5 we give these so-called *compatibility relations* of the irreducible representations of D_3 with the irreducible representations of its two proper subgroups C_2 and C_3.

B.4 Hamilton Operator and Group Theory

In this section we derive the connections between the eigenfunctions and eigenstates of the stationary or time-independent Hamilton operator H and the irreducible representations of a group.

We do not yet specify a particular type of group, but notice that an operator $P(X_i)$ which transforms the system into itself commutes with the Hamilton operator

$$P(X_i)H = HP(X_i), \tag{B.22}$$

where X_i is a symmetry operation of the group. It is known that whenever an operator P commutes with H, there exists a set of functions which are simultaneously eigenfunctions of H and of P. These eigenfunctions are called *basis functions*. The basis functions of the point groups C_{6v}, T_d and O_h are given in the tables of Sect. B.6. We address some properties of basis functions in Sect. B.5. The group of all operations which transforms the system into itself is also known as the *group of* H.

We consider now an eigenvalue E_α of H with eigenfunction(s) $\psi_{\alpha j}$ and with $j = 1, \ldots, m$. If the eigenvalue E_α is not degenerate it follows that $m = 1$. For a degenerate eigenvalue $m > 1$.

Starting with

$$H\psi_{\alpha j} = E_\alpha \psi_{\alpha j} \tag{B.23}$$

we also find

$$P(X_i)H\psi_{\alpha j} = HP(X_i)\psi_{\alpha j} = P(X_i)E_\alpha\psi_{\alpha j} = E_\alpha P(X_i)\psi_{\alpha j}\,, \tag{B.24}$$

since E_α is just a real number. This means that together with $\psi_{\alpha j}$ also $P(X_i)\psi_{\alpha j}$ is an eigenfunction with eigenvalue E_α. Since the $\psi_{\alpha j}$ already include all possible

eigenfunctions with eigenvalue E_α, the $P(X_i)\psi_{\alpha j}$ can only be linear combinations of the $\psi_{\alpha j}$. It follows that

$$P(X_i)\psi_{\alpha k} = \sum_{j=1}^{m} \mathbf{\Gamma}_\alpha(X_i)_{jk}\psi_{\alpha j}\,. \tag{B.25}$$

If we perform this procedure for one symmetry operation X_i and for all $\psi_{\alpha k}$ we get a matrix of coefficients $\mathbf{\Gamma}_\alpha(X_i)_{jk}$. For a non-degenerate eigenvalue this is just a (complex) number of unit magnitude. If we then perform the procedure for all symmetry operations X_i of the group of the Hamiltonian we get a set of matrices $\mathbf{\Gamma}_\alpha(X_i)$. Now comes the crucial point: this set of matrices forms a representation of the group of H, and is in general an irreducible representation.

Due to this fact we can identify the eigenvalue E_α which belongs to the eigenstates $\psi_{\alpha j}$ with the label of the irreducible representation $\mathbf{\Gamma}_\alpha$, i.e., we say that the eigenvalue E_α has symmetry $\mathbf{\Gamma}_\alpha$ or the eigenstates $\psi_{\alpha j}$ of E_α transform according to $\mathbf{\Gamma}_\alpha$.

Since most physical systems have an infinite number of eigenstates but finite groups have only a finite number of non-equivalent irreducible representations, the name of one irreducible representation will occur many times as shown schematically in Fig. B.3. If the representation of the group produced by the eigenfunctions of one eigenenergy E_α is reducible, we speak of an *accidental degeneracy*. An example is shown in Fig. B.3. This means there are actually two different eigenenergies with

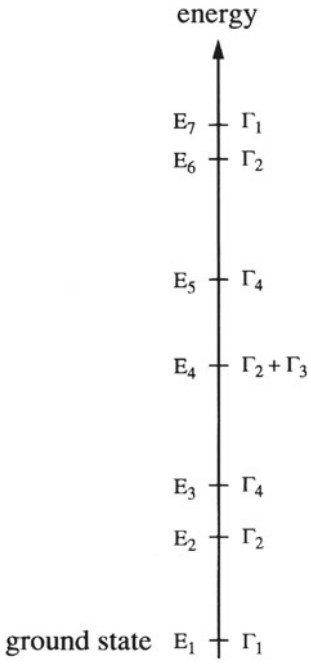


Fig. B.3 The schematic relation of the eigenenergies E_α, and the corresponding irreducible representation $\mathbf{\Gamma}_\alpha$

their sets of eigenfunctions which just happen to be equal for the specific parameters of H, but which could in principle be different.

The important point now is that group theory allows us to make many important predictions, for example, about selection rules or splitting of states under external perturbations, without even knowing the $\mathbf{\Gamma}_\alpha(X_i)_{ij}$ nor the eigenvalues or eigenfunctions explicitly. There is one drawback, however, group theory tells us only whether a certain matrix element is zero or not; it does not say anything about its magnitude if it is non-zero. To getquantitative information we must either use physical arguments to deduce how small or large an effect can be, once group theory has told us that it exists, or we just have to calculate the matrix elements. But here group theory helps us again, by telling us which ones are zero.

Until now we have dealt only with real space. For electronic eigenstates we should also consider spin. Spin does not have a representation in real space. Since we do not want to indulge in a lot of mathematics, we simply note that there exists a group $D_{1/2}$ which describes the behavior of particles with spin 1/2 under point operations (a translation does not affect spin).

To describe electronic wave functions with spin

$$\psi_{\alpha j} = \psi_{\alpha j} \cdot s \,, \tag{B.26}$$

we have to consider the so-called *double group* which is formally just the direct product of $D_{1/2}$ and the corresponding point group $D_{1/2} \otimes$ (point group). This double group is again a group, possessing conjugate elements, classes, irreducible representations, and a character table. We finally end up with more classes and more Γ_α compared to the simple point group without spin.

In the tables given at the end of this chapter for the three most important point groups of semiconductors one can see from the basis functions which ones are the additional Γ_α and classes. For vibrations (phonons) of molecules and of solids spin is not important. Therefore often only the simple and not the double groups are given in corresponding books.

B.5 Applications to Semiconductors Optics

The two groups that describe the symmetry of crystalline solids are, as already mentioned, the translation group and the space group. The translation group is an infinite Abelian group of all translations $\boldsymbol{R}_n$ as introduced in Sects. 9.1 or 14.2. Additionally, the space group contains all other symmetry operations that map the crystal onto itself like inversion, rotations around axes, screw rotations, mirror operations (or reflection) at planes, or glide reflection at planes.

Symmorphic space groups can be decomposed into a semi-direct product of the translation group and a point group that contains symmetry operations that keep at least one point fixed, e.g., inversion, rotation axes or reflection planes. The generators of symmorphic groups do not contain screw rotation axes or reflection glide planes.

Non-symmorphic groups contain such symmetry operations in their generators and cannot be decomposed in the above way. Details about this topic can be found, e.g., in [94H1].

The translation group is obviously Abelian, i.e.,

$$T(\boldsymbol{R}_1)T(\boldsymbol{R}_2) = T(\boldsymbol{R}_1 + \boldsymbol{R}_2) = T(\boldsymbol{R}_2 + \boldsymbol{R}_1) = T(\boldsymbol{R}_2)T(\boldsymbol{R}_1)\,, \tag{B.27}$$

where $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ are two translations of the lattice and T is the translation operator. Consequently every $\boldsymbol{R}_i$, is self-conjugate and forms a class by itself and all irreducible representations are one-dimensional. A detailed inspection of their properties leads to

$$T(\boldsymbol{R}_i)u(\boldsymbol{r}, k) = \mathrm{e}^{\mathrm{i}k\boldsymbol{R}_i}u(\boldsymbol{r}, \boldsymbol{k}) \tag{B.28}$$

since the prefactor must be linear in $\boldsymbol{R}_i$ and of the magnitude one. This short argument shows that the Bloch theorem of Sect. 14.2 also follows from group-theoretical considerations.

Since we have already treated this theorem and its consequences in some detail in Chaps. 9 and 14, we need not further pursue it here, but concentrate now on point groups. We note (again without proof) that according to what we have said above, we can restrict ourselves trivially on symmorphic space groups to the point group.

For non-symmorphic space groups the following statements are correct for the point group we obtain if we replace screw rotation axes by simple rotation axes and glide reflection planes by simple reflection planes [94H1]. In this sense the point group of the wurtzite crystal structure, which contains a screw rotation axis, is $C_{6\mathrm{v}}$ or $6mm$.

The proper consideration of screw rotation axes or glide reflection planes introduces, e.g., phase factors in X-ray scattering but does not change the statements about matrix elements or selection rules given below.

The *point group* of a crystal defined in this sense contains all symmetry operations that transform the crystal into itself and keep at least one point fixed. This group also describes the symmetries of the eigenstates for $\boldsymbol{k} = 0$, i.e., at the $\boldsymbol{\Gamma}$ point in reciprocal space that is the center of the first Brillouin zone.

Since this is also the region of $\boldsymbol{k}$-space that is important for the optical properties of most direct-gap semiconductors, we stick in the following to $\boldsymbol{k} = 0$ (this is the reason why we use $\boldsymbol{\Gamma}$ to label the irreducible representations) and discuss only later what happens when we go to $\boldsymbol{k} \neq 0$. For the moment we note that a statement which is exact for $\boldsymbol{k} = 0$ will be almost correct in the close vicinity of the $\boldsymbol{\Gamma}$-point.

We now focus on the three point groups that are most important for the semiconductors, namely T_d, O_h, and $C_{6\mathrm{v}}$. In Sect. B.6 we give the character tables and the basis functions of these three point groups. The remainder of the 32 point groups can be found in e.g., [63K1, 88L1]. Additionally we give the multiplication tables for the irreducible representations, their compatibility tables with the full rotation group, which is valid for spherical problems like atoms (see below), and with the proper subgroups. There we indicate examples of external perturbations which reduce the symmetry of the point group to that of its subgroups.

The classes with a bar come from the formation of the double group when we include spin and the same is true for the irreducible representations which contain non-integer $\boldsymbol{J}$ and J_z values in their basis functions $\phi(J, J_z)$. Basis functions are examples of functions that possess the symmetry properties of the irreducible representation. Basis functions S_x, S_y and S_z transform like x, y and z, however, without a change of sign under inversion and R means a spherically symmetric function like $x^2 + y^2 + z^2$. A superscript $^+$ or $^-$ appears where inversion C_i is an element of the group and where, consequently, parity is a good quantum number. Obviously $^+$ or $^-$ mean even and odd parity under inversion, respectively.

The information contained in these tables is sufficient to answer many of the questions addressed in the following.

We start with an important rule for the calculation of *transition matrix elements*. If we have to calculate such a matrix element M_{if} (see also Sect. 6.2)

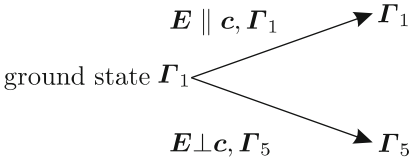
$$M_{if} \propto \langle f\,|H_s|\,i\rangle = \int \psi_f^* H_s \psi_i \mathrm{d}\tau\,, \tag{B.29}$$

where the transition starts from an initial state $|i\rangle$ of symmetry $\boldsymbol{\Gamma}_i$, and ends at a final state $\langle f|$ of symmetry $\boldsymbol{\Gamma}_f$ induced by a perturbation operator H_s that transforms like $\boldsymbol{\Gamma}_s$, then group theory tells us that this matrix element is zero or non-zero according to:

$$\langle f\,|H_\mathrm{s}|\,i\rangle = \begin{cases} \neq 0 & \text{if } \boldsymbol{\Gamma}_f \text{ is contained in the direct product } \boldsymbol{\Gamma}_\mathrm{i} \otimes \boldsymbol{\Gamma}_\mathrm{s} \text{ or} \\ & \text{if } \boldsymbol{\Gamma}_1 \text{ is contained in } \boldsymbol{\Gamma}_f \otimes \boldsymbol{\Gamma}_\mathrm{s} \otimes \boldsymbol{\Gamma}_i\,. \\ & \text{These two statements are equivalent;} \\ 0 & \text{otherwise.} \end{cases} \tag{B.30}$$

This rule allows, for example, to calculate the optical selection rules in crystals. For full spherical symmetry in atoms these rules read, for electric dipole transitions, $\Delta l = \pm 1$; $\Delta m = 0, \pm 1$. For crystals we have to know the symmetry of H_s. It can often be deduced from inspection of the basis functions. The dipole operator transforms in T_d like $\boldsymbol{\Gamma}_5$, in O_h like $\boldsymbol{\Gamma}_5^-$, and in $C_{6\mathrm{v}}$ like $\boldsymbol{\Gamma}_1$ for $\boldsymbol{E} \parallel \boldsymbol{c}$ and like $\boldsymbol{\Gamma}_5$ for $\boldsymbol{E} \perp \boldsymbol{c}$. More information on the symmetries of other perturbations like magnetic fields, uniaxial stress etc. is found e.g., in [63K1, 76C1, 77G1, 85H1, 88L1, 97I1] or in the tables of Sect. B.6.

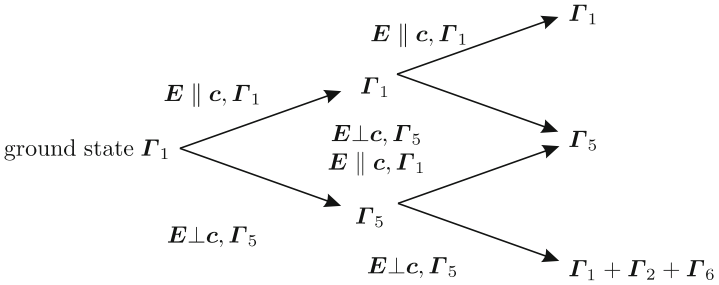
We now use this knowledge to ask which states can be reached by one- and two-photon transitions in materials of $C_{6\mathrm{v}}$ symmetry, starting from the crystal ground state, which has always symmetry $\boldsymbol{\Gamma}_1$. For one-photon transitions, we have to consider matrix elements as in (B.29). Thus for $C_{6\mathrm{v}}$ we find,



For two-photon transitions we have to consider expressions such as:

$$M_{if} \propto \sum_{z_j} \frac{\langle f \left| H^{\mathrm{D},2} \right| z_j \rangle \langle z_j \left| H^{\mathrm{D},1} \right| i \rangle}{\hbar\omega - (E_{z_j} - E_i)} + c.p. \tag{B.31}$$

and find the following selection rules with the help of the multiplication tables



This means that in two-photon absorption states with symmetries $\mathbf{\Gamma}_1$, $\mathbf{\Gamma}_2$, $\mathbf{\Gamma}_5$ and $\mathbf{\Gamma}_6$ can be reached depending on the light polarization.

Perturbation theory is another field in which the vanishing or non-vanishing of a matrix element is essential. We assume that we have an unperturbed stationary Hamiltonian H with non-degenerate eigenstate E_n^0, ψ_n^0. We now apply a small stationary perturbation H_s to get a total Hamiltonian

$$H = H_0 + H_\mathrm{s}\,. \tag{B.32a}$$

The perturbed eigenenergy E_n, and wavefunction ψ_n are then given by

$$E_n = E_n^0 + \langle \psi_n^0 \left| H_\mathrm{s} \right| \psi_n^0 \rangle\,, \tag{B.32b}$$

and

$$\psi_n = \psi_n^0 + \sum_{k \neq n} \frac{\langle \psi_k^0 \left| H_\mathrm{s} \right| \psi_n^0 \rangle \psi_k^0}{E_n^0 - E_k^0}\,. \tag{B.32c}$$

Equation (B.30) tells us now which perturbations can change the eigenenergy and mix the states and which cannot. As stated earlier, group theory does not tell us how large this effect will be.

In the case of a degenerate level E_n^0 with $\psi_{n,i}^0$, perturbation theory leads to a secular equation given by the determinant

$$\det \left| \langle \psi_{n,i}^0 \left| H_\mathrm{s} \right| \psi_{n,j}^0 \rangle - E\delta_{ij} \right| = 0\,, \tag{B.33}$$

and again (B.30) can predict which matrix elements vanish. The problem of degenerate states has been developed even further with a technique known as *invariant*

expansion. A discussion of this technique is beyond the scope of this book but the reader may find an elaborate presentation in [76C1, 77G1, 85H1].

Partly equivalent information is obtained if we consider that the perturbation reduces the symmetry of the problem, so that the perturbed problem corresponds to a subgroup of the unperturbed one. If the irreducible representation of E_n^0 is still irreducible in the subgroup, E_n may shift but does not split. If it decays into more than one irreducible representation in the subgroup, E_n^0 can split into a corresponding number of different levels under the influence of H_s. If some of the new eigenvalues still coincide we have an accidental degeneracy.

To illustrate this fact, inspection of the compatibility relations of C_{6v} shows that an electrostatic field $\boldsymbol{E} \perp \boldsymbol{c}$ ($\boldsymbol{\Gamma}_5$) will result in a splitting of a $\boldsymbol{\Gamma}_5$ level into two states of symmetries $\boldsymbol{\Gamma}_1$ and $\boldsymbol{\Gamma}_2$ of the group C_s.

Another case in which group theory is useful concerns *product wavefunctions*. If, for example, we know that the spatial part of the hole wavefunction $\phi_h(\boldsymbol{r}_n)$ in T_d transforms like $\boldsymbol{\Gamma}_5$ (this results from an atomic p-state as will be shown below) and we want to add spin σ we get

$$\psi = \phi_n^{\boldsymbol{\Gamma}_5}(\boldsymbol{r}_h) \cdot \boldsymbol{s} \,. \tag{B.34}$$

In order to uncover the symmetry of the total wave function we just form the direct product of the symmetry of the spatial part (here $\boldsymbol{\Gamma}_5$) and of the irreducible representation of spin 1/2. For T_d this reads

$$\boldsymbol{\Gamma}_{\text{total}} = \boldsymbol{\Gamma}_{5,\text{space}} \otimes \boldsymbol{\Gamma}_{1/2} = \boldsymbol{\Gamma}_5 \otimes \boldsymbol{\Gamma}_6 = \boldsymbol{\Gamma}_7 \oplus \boldsymbol{\Gamma}_8 \,. \tag{B.35}$$

This means that the spin results in a splitting into two states of symmetries $\boldsymbol{\Gamma}_7$ and $\boldsymbol{\Gamma}_8$. From physical arguments we know that this is the spin-orbit splitting.

A very important example of product wavefunctions in semiconductor optics are exciton wavefunctions. As shown in Sect. 20.2, they are a product of the electron and hole wavefunctions and of the envelope function. The possible symmetries of an exciton thus result from the direct product of the symmetries of the electron, hole and envelope functions, i.e.,

$$\boldsymbol{\Gamma}_{\text{exciton}} = \boldsymbol{\Gamma}_{\text{el}} \otimes \boldsymbol{\Gamma}_{\text{h}} \otimes \boldsymbol{\Gamma}_{\text{env}} \,. \tag{B.36}$$

For the ground state (main quantum number $n_B = 1$) $\boldsymbol{\Gamma}_{\text{env}}$ is always $\boldsymbol{\Gamma}_1$. So in T_d, for example, we find excitons formed with the hole in the $\boldsymbol{\Gamma}_7$ or $\boldsymbol{\Gamma}_8$ valence band and the electron in the $\boldsymbol{\Gamma}_6$ conduction band

$$\boldsymbol{\Gamma}_6 \otimes \boldsymbol{\Gamma}_8 \otimes \boldsymbol{\Gamma}_1 = \boldsymbol{\Gamma}_3 \oplus \boldsymbol{\Gamma}_4 \oplus \boldsymbol{\Gamma}_5 \tag{B.37a}$$

$$\boldsymbol{\Gamma}_6 \otimes \boldsymbol{\Gamma}_7 \otimes \boldsymbol{\Gamma}_1 = \boldsymbol{\Gamma}_2 \oplus \boldsymbol{\Gamma}_5 \,. \tag{B.37b}$$

The $\boldsymbol{\Gamma}_5$ is the spin singlet state, which can be reached from the ground state by an electric-dipole transition; the $\boldsymbol{\Gamma}_3$, $\boldsymbol{\Gamma}_4$ and $\boldsymbol{\Gamma}_2$ states are dipole forbidden. Actually they

are the triplet states with parallel electron and hole spins. The number of possible exciton states increases rapidly for $n_B > 1$.

The reduction of symmetry in a quantum well compared to the parent bulk materials is the next important application we want to discuss. In Table B.6 we give the irreducible representations, the compatibility relations and the selection rules when going from the symmetry of a bulk zincblende structure to one of a QW grown on a (001) surface.

A single ideal interface between two different materials both having the point group T_d reduces the symmetry to C_{2v} [96I1, 97I1]. A quantum well, i.e., two close-lying interfaces, results in symmetry D_{2d} if the two materials have a common anion or cation, e.g., AB/AC. This situation is treated in Table B.6. In the case of a well/barrier combination without a common anion or cation AB/CD, the total symmetry may be D_{2d} or C_{2v} depending on the termination of the interfaces. For C_{2v} symmetry, the orthogonal directions [110] and $[1\bar{1}0]$ in the plane of the quantum well are no longer equivalent. This may result in polarization anisotropies of the optical properties or in different band offsets at both interfaces and in internal electric fields if, e.g., the dipole layers on both sides are different. For more details see, e.g., [97I1, 00S1, 02M1, 03T1]).

Going one step further, we come to the more complicated entities such as bound-exciton complexes or biexcitons. The procedure is basically as above, i.e.,

Table B.6 Group theoretical relations and selection rules for zincblende-type bulk material and quantum wells grown on (100) oriented substrates [93J1]

	Point group T_d (GaAs, bulk)	Point group D_{2d} (GaAs QW on (001) oriented substrate)	Comments
Irreducible representation of lowest conduction band	Γ_6	Γ_6	
Irreducible representation of highest valence band (does not apply to CuCl or a few other exceptions)	Γ_8	Γ_6 Γ_7	Heavy holes Light holes
Excitons with S, i.e., Γ_1 envelope function	$\Gamma_6^{LB} \otimes \Gamma_8^{VB} \otimes \Gamma_1^{env}$ $= \Gamma_3 + \Gamma_4 + \Gamma_5$	$\Gamma_6^{LB} \otimes \Gamma_6^{VB} \otimes \Gamma_1^{env}$ $= \Gamma_1 + \Gamma_2 + \Gamma_5$ $\Gamma_6^{LB} \otimes \Gamma_7^{VB} \otimes \Gamma_1^{env}$ $= \Gamma_3 + \Gamma_4 + \Gamma_5$	Heavy-hole excitons Light-hole excitons
Dipole-allowed transitions	Γ_5	Γ_5 for $\boldsymbol{E} \perp z$ Γ_4 for $\boldsymbol{E} \parallel z$	z is the growth direction, i.e., the normal to the QW plane

$$\mathbf{\Gamma}_{\text{biex}} = (\mathbf{\Gamma}_{\text{el}} \otimes \mathbf{\Gamma}_{\text{el}})^{\pm} \otimes (\mathbf{\Gamma}_{\text{h}} \otimes \mathbf{\Gamma}_{\text{h}})^{\pm} \mathbf{\Gamma}_{\text{env}}^{\mp\mp} \,. \tag{B.38}$$

However, one now has to consider that these systems contain partly indistinguishable fermions. Therefore the total wave-function must change sign under the exchange of two identical particles. This means that if the combination of the two electrons $(\mathbf{\Gamma}_{\text{el}} \otimes \mathbf{\Gamma}_{\text{el}})^{-}$ changes sign under exchange, then the envelope function must have even parity under this operation and vice versa. The same holds for the holes.

The parity of the combinations can be deduced from the coupling coefficients in tables like [63K1]. We do not go into details here but mention that the ground state of the biexciton always has an envelope of $\mathbf{\Gamma}_1$ symmetry. The possible combinations for the excitons of (B.37a) and (B.37b) are in the ground state $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_1 \oplus \mathbf{\Gamma}_3 \oplus \mathbf{\Gamma}_5$ [77G1, 85H1], respectively. In C_{6v} the biexciton containing two holes from the same valence band always has only symmetry $\mathbf{\Gamma}_1$ in the ground state, but with a hole from the $\mathbf{\Gamma}_9$ and another from a $\mathbf{\Gamma}_7$ valence band it is possible to construct in addition $\mathbf{\Gamma}_5$ and $\mathbf{\Gamma}_6$ biexciton levels [77G1, 85H1].

The bands in semiconductors often still contain some information about the parent atomic orbitals, especially in the lattice periodic part $u_k(\boldsymbol{r})$ of the Bloch function. The atomic orbitals result from a spherically symmetric problem. Therefore it is reasonable to shortly describe this group and its compatibility relations with the point groups of semiconductors.

In the full *spherical rotation group*, the system can be rotated around any axis by any angle and is transformed into itself. It can be shown that all rotations by the same angle ϕ but around arbitrary axes are in the same class. The full rotation group can therefore be considered as a *continuous group* (Fig. B.4) with rotation angle $0 \leq \phi < 2\pi$. Consequently there must be a correspondingly infinite number of irreducible representations $\mathbf{\Gamma}_l$. Figure B.4 gives the character of the representation. For the identity element E we get the dimensionality of $\mathbf{\Gamma}_l$ which is obviously $2l + 1$. The basis functions are the spherical harmonics $Y_l^m(\phi, \theta)$.

The compatibility relations between the full rotation group for even and odd parity, called $D_{L_z}^{\pm}$ and $D_{J_z}^{\pm}$ and the three point groups T_d, O_{h} and C_{6v} are given in Sect. B.6. These tables may be used as follows: In a tight-binding approximation we would assume that the uppermost valence band of CdS is formed from the filled $3p$ levels of S^{2-}, while the lowest conduction band comes from the empty $5s$ levels of Cd^{2+}. From the compatibility tables we learn that this results in the following band symmetries for C_{6v}, at the $\mathbf{\Gamma}$-point:

$$\begin{aligned} &\text{CB}: \quad 5s \text{ levels} \Rightarrow D_{1/2} \Rightarrow \mathbf{\Gamma}_7 \,; \\ &\text{VB}: \quad 3p \text{ levels} \Rightarrow \begin{cases} D_{3/2} \Rightarrow \mathbf{\Gamma}_7 + \mathbf{\Gamma}_9 \\ D_{1/2} \Rightarrow \mathbf{\Gamma}_7 \,. \end{cases} \end{aligned} \tag{B.39}$$

For T_d symmetry one finds similarly

$$\text{CB}: \ \mathbf{\Gamma}_6, \quad \text{VB}: \ \mathbf{\Gamma}_7 + \mathbf{\Gamma}_8 \,. \tag{B.40}$$

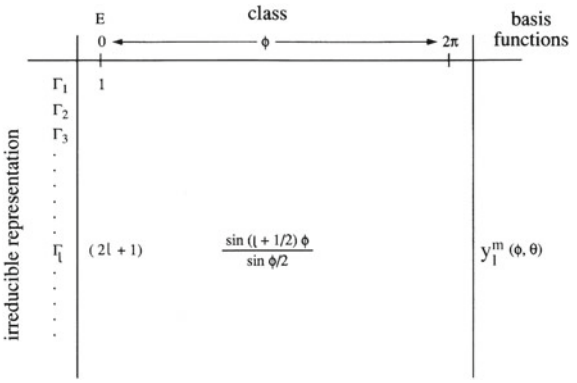


Fig. B.4 The character table of the spherical rotation group

This implies a twofold splitting in the valence band of T_d symmetry, which comes from the spin-orbit interaction as already mentioned, and a threefold splitting in C_{6v} again arising from spin-orbit splitting and, in addition, from the hexagonal crystal field.

Note, that for covalent sp^3 binding the top of the valence band and the bottom of the conduction band also have p- and s-character, thus leading to the same results as for ionic binding.

Group theory cannot tell us the magnitude of the splittings nor the ordering of the bands. It is found, however, that for T_d symmetry the $\mathbf{\Gamma}_8$ state usually forms the upper valence band and the $\mathbf{\Gamma}_7$ comes below (except in CuCl which has the reverse ordering). In C_{6v} one usually has an uppermost valence band of symmetry $\mathbf{\Gamma}_9$ and two $\mathbf{\Gamma}_7$ levels below (except for ZnO, where the two upper bands are also most probably interchanged). See Sects. 15.2, 21.1, and 24.1

It is clear that angular momentum is no longer a good *quantum number* in solids. The good quantities are the $\mathbf{\Gamma}_i$. With great care, however, one can sometimes still use arguments based on angular momentum. In the wurtzite structure the $\boldsymbol{c}$ axis is the quantization axis and J_z is to some extent usable for discussions. The p-orbitals forming the valence band have $l = 1$ and $m = 0, \pm 1$. The $m = 0$ states have symmetry $\mathbf{\Gamma}_1$ (or $\mathbf{\Gamma}_2$) the $m = \pm 1$ states have symmetry $\mathbf{\Gamma}_5$. With the z-component of the spin of $\pm 1/2$ we can then produce $J_z = \pm\hbar/2$ and $J_z = \pm 3\hbar/2$ states corresponding to $\mathbf{\Gamma}_7$ and $\mathbf{\Gamma}_9$, respectively. For values of J_z larger than $3\hbar/2$, however, the rather shaky approach above breaks down completely. See also the basis functions in the tables for O_h and T_d.

We want to conclude this section by a look at the symmetry of the semiconductor bandstructure away from points of high symmetry.

Let us consider a state in the band structure away from the $\mathbf{\Gamma}$ point, i.e., with $\boldsymbol{k} \neq 0$. First we apply all symmetry operations of the point group to its $\boldsymbol{k}$-vector as shown schematically in Fig. B.5a for a two-dimensional square lattice. The result is a number of $\boldsymbol{k}$-vectors, which are known as the *star of* $\boldsymbol{k}$. If $\boldsymbol{k}$ is in a "general" position there are as many $\boldsymbol{k}$-vectors in the star as the group has elements. On the other

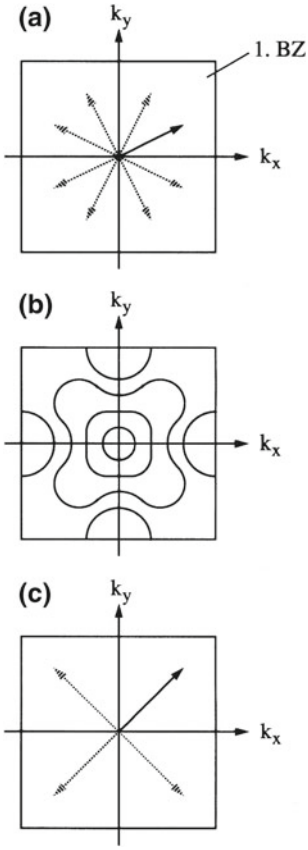


Fig. B.5 The star of $\boldsymbol{k}$ (**a**), the qualitative shape of surfaces of constant energy $E(\boldsymbol{k})$ (**b**) and a $\boldsymbol{k}$-vector of higher symmetry (**c**)

hand, all these symmetry operations transform the system into itself. Consequently the properties must be the same for all elements of the star of $\boldsymbol{k}$. We illustrate this statement in Fig. B.5b for the dispersion $E(\boldsymbol{k})$. We give contours of constant energy and see clearly that the star of $\boldsymbol{k}$ allows a fourfold band warping as indeed occurs for T_d and O_h symmetries, simply because cubic symmetry is lower than spherical symmetry.

If $\boldsymbol{k}$ is not in a "general" direction but coincides with one of the symmetry lines or planes of the system, then several of the elements of the star of $\boldsymbol{k}$ coincide as shown in Fig. B.5c. The symmetry operations that transform $\boldsymbol{k}$ into itself are known as the *group of* $\boldsymbol{k}$. This group of $\boldsymbol{k}$ is a subgroup of the full point group. The irreducible representations at the $\boldsymbol{\Gamma}$-point are consequently also representations of the subgroup. If we start at the $\boldsymbol{\Gamma}$-point with a certain energy state which transforms like $\boldsymbol{\Gamma}_\alpha$ and move along some symmetry direction, the energy level $\boldsymbol{\Gamma}_\alpha$ may or may not split, depending on whether $\boldsymbol{\Gamma}_\alpha$ becomes a reducible representation or remains an irreducible one in the (sub) group of $\boldsymbol{k}$.

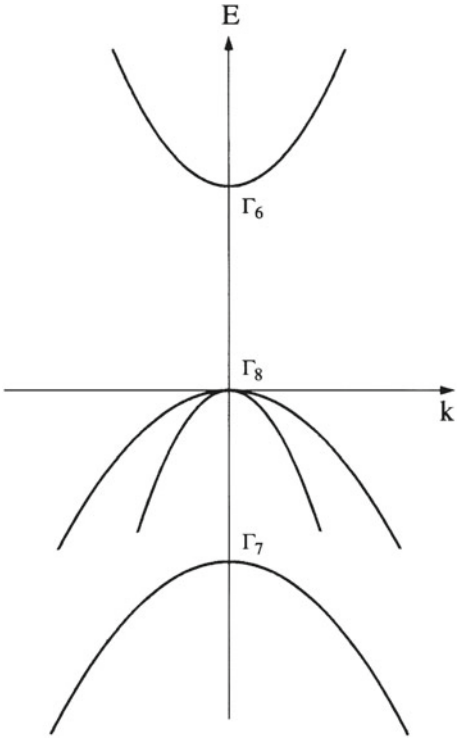


Fig. B.6 The usual bandstructure around the $\mathbf{\Gamma}$-point for T_d symmetry (Compare with Fig. 15.9)

A typical example is the splitting of the fourfold degenerate $\mathbf{\Gamma}_8$ state in T_d symmetry for $\boldsymbol{k} \neq 0$ into two twofold degenerate states known as heavy and light-hole bands (Figs. 15.9 and B.6), where the states $(J, J_z) = (3/2, \pm 3/2)$ correspond to the heavy-hole valence band and $(3/2, \pm 1/2)$ corresponds to the light-hole valence band (see e.g., [74B1, 03T1]). The $\mathbf{\Gamma}_6$ and $\mathbf{\Gamma}_7$ states do not show such behavior.

For a general orientation of the $\boldsymbol{k}$-vector all degeneracies can be lifted. In bandstructure theory one thus usually calculates $E(\boldsymbol{k})$ in directions of high symmetry (see Sect. 15.2) to exploit the symmetry properties described by group theory and tries to extrapolate in other directions if necessary.

The final symmetry operation that we address is the invariance of a microscopic physical system like a crystalline semiconductor under *time reversal*. If we neglect spin for the moment, invariance under time reversal has the following consequence for dispersion relations:

$$E(\boldsymbol{k}) = E(-\boldsymbol{k})\,, \tag{B.41}$$

even if the point group does not include the inversion. This phenomenon is known as *Kramers degeneracy*.

In a power series expansion of the dispersion

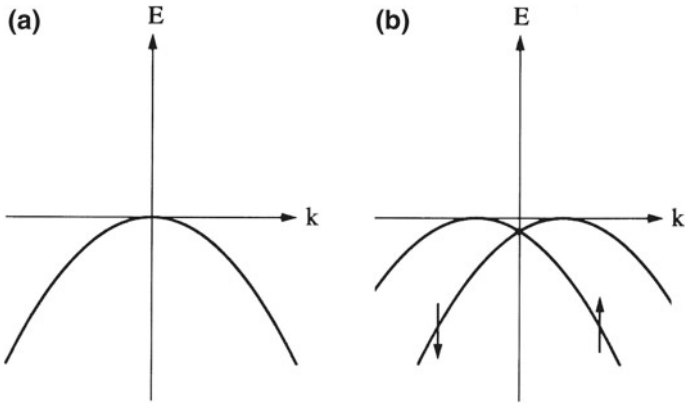


Fig. B.7 Schematic drawing of the Kramers degeneracy without spin (**a**) and with spin (**b**) as occurs e.g., for Γ_7 bands in C_{6v} symmetry for $\boldsymbol{k} \perp \boldsymbol{c}$

$$E(\boldsymbol{k}) = \sum_{n=0}^{\infty} a_n \boldsymbol{k}^n \tag{B.42}$$

this permits only even values for n (Fig. B.7a). If we now include the spin, (B.41) changes to

$$E(\boldsymbol{k}, \uparrow) = E(-\boldsymbol{k}, \downarrow)\,. \tag{B.43}$$

This means that time reversal flips the spin. This condition can also be fulfilled with odd powers in the expansion (B.42) and, in particular, allows terms linear in $\boldsymbol{k}$ (Fig. B.7b).

Detailed group-theoretical investigations show that *k-linear terms* are possible for Γ_8 states in T_d symmetry or for Γ_7 states $\boldsymbol{k} \perp \boldsymbol{c}$ in C_{6v} symmetry, but not for $\boldsymbol{k} \parallel \boldsymbol{c}$ and not for $\boldsymbol{\Gamma}_9$ levels.

Indeed such $\boldsymbol{k}$-linear terms are known; for example, for the B-exciton in CdS which contains a hole from the Γ_7 valence band or for the A excitons in ZnO. See Sect. 21.1.

B.6 Some Selected Group Tables

In this section, we give some tables for the three most important groups of semiconductors. We start in Table B.7 with the character tables of the point group C_{6v}. In the first column we give the names of the irreducible representations in various notations, in the last one typical basis functions. From them it becomes clear, which representation belongs to the simple and which one to the double group including spin. These are here $\boldsymbol{\Gamma}_1$–$\boldsymbol{\Gamma}_6$ and $\boldsymbol{\Gamma}_7$–$\boldsymbol{\Gamma}_9$, respectively. The next tables give the multiplication tables of the irreducible representations and the compatibility relation with the subgroups as well with the full rotation group (Tables B.8, B.9 and B.10).

Tables B.11, B.12, B.13 and B.14 give the same information for T_d. For O_h we give only two tables for characters and the compatibility relation (Tables B.15 and B.16).

Table B.7 Character table and basis functions of the point group C_{6v} (or $6mm$). Below E_2 or Γ_6 are the additional irreducible representations of the double group which includes spin

C_{6v}		E	$\bar{E}$	C_2	$2C_3$	$2\bar{C}_3$	$2C_6$	$2\bar{C}_6$	$3\bar{\sigma}_d$	$3\bar{\sigma}_v$	Time reversal	Basis functions
A_1	Γ_1	1	1	1	1	1	1	1	1	1	a	R or z
A_2	Γ_2	1	1	1	1	1	1	1	−1	−1	a	S_z
B_1	Γ_3	1	1	−1	1	1	−1	−1	1	−1	a	$x^3 - 3xy^2$
B_2	Γ_4	1	1	−1	1	1	−1	−1	−1	1	a	$x^3 - 3xy^2$
E_1	Γ_5	2	2	−2	−1	−1	1	1	0	0	a	$(S_x - \mathrm{i}S_y)$, $-(S_x + \mathrm{i}S_y)$
E_2	Γ_6	2	2	2	−1	−1	−1	−1	0	0	a	$\Gamma_3 \otimes \Gamma_5$
	Γ_7	2	−2	0	1	−1	$\sqrt{3}$	$-\sqrt{3}$	0	0	c	$\phi(1/2, -1/2)$, $\phi(1/2, 1/2)$
	Γ_8	2	−2	0	1	−1	$-\sqrt{3}$	$\sqrt{3}$	0	0	c	$\Gamma_7 \otimes \Gamma_3$
	Γ_9	2	−2	0	−2	2	0	0	0	0	c	$\phi(3/2, -3/2)$, $\phi(3/2, 3/2)$

Table B.8 Multiplication table for the irreducible representations of the group $C_{6\mathrm{v}}$

Γ_1	Γ_2	Γ_3	Γ_4	Γ_5	Γ_6	Γ_7	Γ_8	Γ_9	
Γ_1	Γ_2	Γ_3	Γ_4	Γ_5	Γ_6	Γ_7	Γ_8	Γ_9	Γ_1
	Γ_1	Γ_4	Γ_3	Γ_5	Γ_6	Γ_7	Γ_8	Γ_9	Γ_2
		Γ_1	Γ_2	Γ_6	Γ_5	Γ_8	Γ_7	Γ_9	Γ_3
			Γ_1	Γ_6	Γ_5	Γ_8	Γ_7	Γ_9	Γ_4
				$\Gamma_1 + \Gamma_2 + \Gamma_6$	$\Gamma_3 + \Gamma_4 + \Gamma_5$	$\Gamma_7 + \Gamma_9$	$\Gamma_8 + \Gamma_9$	$\Gamma_7 + \Gamma_8$	Γ_5
					$\Gamma_1 + \Gamma_2 + \Gamma_6$	$\Gamma_8 + \Gamma_9$	$\Gamma_7 + \Gamma_9$	$\Gamma_7 + \Gamma_8$	Γ_6
						$\Gamma_1 + \Gamma_2 + \Gamma_5$	$\Gamma_3 + \Gamma_4 + \Gamma_6$	$\Gamma_5 + \Gamma_6$	Γ_7
							$\Gamma_1 + \Gamma_2 + \Gamma_5$	$\Gamma_5 + \Gamma_6$	Γ_8
								$\Gamma_1 + \Gamma_2$ $\Gamma_3 + \Gamma_4$	Γ_9

Table B.9 Compatibility table of the irreducible representations of the group C_{6v} with the irreducible representations of its proper subgroups

$C_{6v} : E(z)$	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_4$	$\mathbf{\Gamma}_5$	$\mathbf{\Gamma}_6$	$\mathbf{\Gamma}_7$	$\mathbf{\Gamma}_8$	$\mathbf{\Gamma}_9$
$C_6 : H(z)$	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_4$	$\mathbf{\Gamma}_4$	$\mathbf{\Gamma}_5 + \mathbf{\Gamma}_6$	$\mathbf{\Gamma}_2 + \mathbf{\Gamma}_3$	$\mathbf{\Gamma}_7 + \mathbf{\Gamma}_8$	$\mathbf{\Gamma}_9 + \mathbf{\Gamma}_{10}$	$\mathbf{\Gamma}_{11} + \mathbf{\Gamma}_{12}$
C_{3v}	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_4$	$\mathbf{\Gamma}_4$	$\mathbf{\Gamma}_5 + \mathbf{\Gamma}_6$
C_{2v}	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_4$	$\mathbf{\Gamma}_2 + \mathbf{\Gamma}_4$	$\mathbf{\Gamma}_1 + \mathbf{\Gamma}_3$	$\mathbf{\Gamma}_5$	$\mathbf{\Gamma}_5$	$\mathbf{\Gamma}_5$
$C_s : E(x) : H(y)$	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_1 + \mathbf{\Gamma}_2$	$\mathbf{\Gamma}_1 + \mathbf{\Gamma}_2$	$\mathbf{\Gamma}_3 + \mathbf{\Gamma}_4$	$\mathbf{\Gamma}_3 + \mathbf{\Gamma}_4$	$\mathbf{\Gamma}_3 + \mathbf{\Gamma}_4$
$C_s : H(x) : E(y)$	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_1 + \mathbf{\Gamma}_2$	$\mathbf{\Gamma}_1 + \mathbf{\Gamma}_2$	$\mathbf{\Gamma}_3 + \mathbf{\Gamma}_4$	$\mathbf{\Gamma}_3 + \mathbf{\Gamma}_4$	$\mathbf{\Gamma}_3 + \mathbf{\Gamma}_4$

Table B.10 Compatibility table of the irreducible representations of the full rotation group of angular momentum $L_z \leq 6$ and $J_z \leq \frac{13}{2}$ with the irreducible representations of the group C_{6v}

D_0^+	$\mathbf{\Gamma}_1$	D_0^-	$\mathbf{\Gamma}_2$
D_1^+	$\mathbf{\Gamma}_2 + \mathbf{\Gamma}_5$	D_1^-	$\mathbf{\Gamma}_1 + \mathbf{\Gamma}_5$
D_2^+	$\mathbf{\Gamma}_1 + \mathbf{\Gamma}_5 + \mathbf{\Gamma}_6$	D_2^-	$\mathbf{\Gamma}_2 + \mathbf{\Gamma}_5 + \mathbf{\Gamma}_6$
D_3^+	$\mathbf{\Gamma}_2 + \mathbf{\Gamma}_3 + \mathbf{\Gamma}_4 + \mathbf{\Gamma}_5 + \mathbf{\Gamma}_6$	D_3^-	$\mathbf{\Gamma}_1 + \mathbf{\Gamma}_3 + \mathbf{\Gamma}_4 + \mathbf{\Gamma}_5 + \mathbf{\Gamma}_6$
D_4^+	$\mathbf{\Gamma}_1 + \mathbf{\Gamma}_3 + \mathbf{\Gamma}_4 + \mathbf{\Gamma}_5 + 2\mathbf{\Gamma}_6$	D_4^-	$\mathbf{\Gamma}_2 + \mathbf{\Gamma}_3 + \mathbf{\Gamma}_4 + \mathbf{\Gamma}_5 + 2\mathbf{\Gamma}_6$
D_5^+	$\mathbf{\Gamma}_2 + \mathbf{\Gamma}_3 + \mathbf{\Gamma}_4 + 2\mathbf{\Gamma}_5 + 2\mathbf{\Gamma}_6$	D_5^-	$\mathbf{\Gamma}_1 + \mathbf{\Gamma}_3 + \mathbf{\Gamma}_4 + 2\mathbf{\Gamma}_5 + 2\mathbf{\Gamma}_6$
D_6^+	$2\mathbf{\Gamma}_1 + \mathbf{\Gamma}_2 + \mathbf{\Gamma}_3 + \mathbf{\Gamma}_4 + 2\mathbf{\Gamma}_5 + 2\mathbf{\Gamma}_6$	D_6^-	$\mathbf{\Gamma}_1 + 2\mathbf{\Gamma}_2 + \mathbf{\Gamma}_3 + \mathbf{\Gamma}_4 + 2\mathbf{\Gamma}_5 + 2\mathbf{\Gamma}_6$
$D_{1/2}^+$	$\mathbf{\Gamma}_7$	$D_{1/2}^-$	$\mathbf{\Gamma}_7$
$D_{3/2}^+$	$\mathbf{\Gamma}_7 + \mathbf{\Gamma}_9$	$D_{3/2}^-$	$\mathbf{\Gamma}_7 + \mathbf{\Gamma}_9$
$D_{5/2}^+$	$\mathbf{\Gamma}_7 + \mathbf{\Gamma}_8 + \mathbf{\Gamma}_9$	$D_{5/2}^-$	$\mathbf{\Gamma}_7 + \mathbf{\Gamma}_8 + \mathbf{\Gamma}_9$
$D_{7/2}^+$	$\mathbf{\Gamma}_7 + 2\mathbf{\Gamma}_8 + 2\mathbf{\Gamma}_9$	$D_{7/2}^-$	$\mathbf{\Gamma}_7 + 2\mathbf{\Gamma}_8 + 2\mathbf{\Gamma}_9$
$D_{9/2}^+$	$\mathbf{\Gamma}_7 + 2\mathbf{\Gamma}_8 + \mathbf{\Gamma}_9$	$D_{9/2}^-$	$\mathbf{\Gamma}_7 + 2\mathbf{\Gamma}_8 + 2\mathbf{\Gamma}_9$
$D_{11/2}^+$	$2\mathbf{\Gamma}_7 + 2\mathbf{\Gamma}_8 + 2\mathbf{\Gamma}_9$	$D_{11/2}^-$	$2\mathbf{\Gamma}_7 + 2\mathbf{\Gamma}_8 + 2\mathbf{\Gamma}_9$
$D_{13/2}^+$	$3\mathbf{\Gamma}_7 + 2\mathbf{\Gamma}_8 + 2\mathbf{\Gamma}_9$	$D_{13/2}^-$	$3\mathbf{\Gamma}_7 + 2\mathbf{\Gamma}_8 + 2\mathbf{\Gamma}_9$

Table B.11 Character table and basis functions for the irreducible representations of the point group T_d (or $\bar{4}3\,m$)

T_d		E	$\bar{E}$	$8C_3$	$8\bar{C}_3$	$3C_2$	$6S_4$	$6\bar{S}_4$	$6\sigma_d$	Time reversal	Basis functions
A_1	$\mathbf{\Gamma}_1$	1	1	1	1	1	1	1	1	a	R or xyz
A_2	$\mathbf{\Gamma}_2$	1	1	1	1	1	-1	-1	-1	a	$S_x S_y S_z$
E	$\mathbf{\Gamma}_3$	2	2	-1	-1	2	0	0	0	a	$(2z^2 - x^2 - y^2)$ $\sqrt{3}(s^2 - y^2)$
F_1	$\mathbf{\Gamma}_4$	3	3	0	0	-1	1	1	-1	a	S_x, S_y, S_z
F_2	$\mathbf{\Gamma}_5$	3	3	0	0	-1	-1	-1	1	a	x, y, z
$\bar{E}_1$	$\mathbf{\Gamma}_6$	2	-2	1	-1	0	$\sqrt{2}$	$-\sqrt{2}$	0	c	$\phi(1/2, -1/2)$ $\phi(1/2, 1/2)$
$\bar{E}_2$	$\mathbf{\Gamma}_7$	2	-2	1	-1	0	$-\sqrt{2}$	$\sqrt{2}$	0	c	$\mathbf{\Gamma}_6 \times \mathbf{\Gamma}_2$
$\bar{E}_3$	$\mathbf{\Gamma}_8$	4	-4	-1	1	0	0	0	0	c	$\phi(3/2, -3/2)$ $\phi(3/2, -1/2)$ $\phi(3/2, 1/2)$ $\phi(3/2, 3/2)$

Table B.12 Multiplication table of the irreducible representations of the group T_d

$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_4$	$\mathbf{\Gamma}_5$	$\mathbf{\Gamma}_6$	$\mathbf{\Gamma}_7$	$\mathbf{\Gamma}_8$	
$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_4$	$\mathbf{\Gamma}_5$	$\mathbf{\Gamma}_6$	$\mathbf{\Gamma}_7$	$\mathbf{\Gamma}_8$	$\mathbf{\Gamma}_1$
	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_5$	$\mathbf{\Gamma}_4$	$\mathbf{\Gamma}_7$	$\mathbf{\Gamma}_6$	$\mathbf{\Gamma}_8$	$\mathbf{\Gamma}_2$
		$\mathbf{\Gamma}_1 + \mathbf{\Gamma}_2 + \mathbf{\Gamma}_3$	$\mathbf{\Gamma}_4 + \mathbf{\Gamma}_5$	$\mathbf{\Gamma}_4 + \mathbf{\Gamma}_5$	$\mathbf{\Gamma}_8$	$\mathbf{\Gamma}_8$	$\mathbf{\Gamma}_6 + \mathbf{\Gamma}_7 + \mathbf{\Gamma}_8$	$\mathbf{\Gamma}_3$
			$\mathbf{\Gamma}_1 + \mathbf{\Gamma}_3 + \mathbf{\Gamma}_4 + \mathbf{\Gamma}_5$	$\mathbf{\Gamma}_2 + \mathbf{\Gamma}_3 + \mathbf{\Gamma}_4 + \mathbf{\Gamma}_5$	$\mathbf{\Gamma}_6 + \mathbf{\Gamma}_8$	$\mathbf{\Gamma}_7 + \mathbf{\Gamma}_8$	$\mathbf{\Gamma}_6 + \mathbf{\Gamma}_7 + 2\mathbf{\Gamma}_8$	$\mathbf{\Gamma}_4$
				$\mathbf{\Gamma}_1 + \mathbf{\Gamma}_3 + \mathbf{\Gamma}_4 + \mathbf{\Gamma}_5$	$\mathbf{\Gamma}_7 + \mathbf{\Gamma}_8$	$\mathbf{\Gamma}_6 + \mathbf{\Gamma}_8$	$\mathbf{\Gamma}_6 + \mathbf{\Gamma}_7 + 2\mathbf{\Gamma}_8$	$\mathbf{\Gamma}_5$
					$\mathbf{\Gamma}_1 + \mathbf{\Gamma}_4$	$\mathbf{\Gamma}_2 + \mathbf{\Gamma}_5$	$\mathbf{\Gamma}_3 + \mathbf{\Gamma}_4 + \mathbf{\Gamma}_5$	$\mathbf{\Gamma}_6$
						$\mathbf{\Gamma}_1 + \mathbf{\Gamma}_4$	$\mathbf{\Gamma}_3 + \mathbf{\Gamma}_4 + \mathbf{\Gamma}_5$	$\mathbf{\Gamma}_7$
							$\mathbf{\Gamma}_1 + \mathbf{\Gamma}_2 + \mathbf{\Gamma}_3 + 2\mathbf{\Gamma}_4 + 2\mathbf{\Gamma}_5$	$\mathbf{\Gamma}_8$

Table B.13 Compatibility table of the irreducible representations of T_d with the irreducible representations of its proper subgroups

T_d	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_4$	$\mathbf{\Gamma}_5$	$\mathbf{\Gamma}_6$	$\mathbf{\Gamma}_7$	$\mathbf{\Gamma}_8$
T	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_2+\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_4$	$\mathbf{\Gamma}_4$	$\mathbf{\Gamma}_5$	$\mathbf{\Gamma}_5$	$\mathbf{\Gamma}_6+\mathbf{\Gamma}_7$
D_{2d}	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_1+\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_2+\mathbf{\Gamma}_5$	$\mathbf{\Gamma}_4+\mathbf{\Gamma}_5$	$\mathbf{\Gamma}_6$	$\mathbf{\Gamma}_7$	$\mathbf{\Gamma}_6+\mathbf{\Gamma}_7$
$C_{3\mathrm{v}}: E(w)$	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_2+\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_1+\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_4$	$\mathbf{\Gamma}_4$	$\mathbf{\Gamma}_4+\mathbf{\Gamma}_5+\mathbf{\Gamma}_6$
$S_1: H(z)$	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_1+\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_1+\mathbf{\Gamma}_3+\mathbf{\Gamma}_4$	$\mathbf{\Gamma}_2+\mathbf{\Gamma}_3+\mathbf{\Gamma}_4$	$\mathbf{\Gamma}_5+\mathbf{\Gamma}_6$	$\mathbf{\Gamma}_7+\mathbf{\Gamma}_8$	$\mathbf{\Gamma}_5+\mathbf{\Gamma}_6+\mathbf{\Gamma}_7+\mathbf{\Gamma}_5$
$C_3: H(w)$	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_2+\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_1+\mathbf{\Gamma}_2+\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_1+\mathbf{\Gamma}_2+\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_4+\mathbf{\Gamma}_5$	$\mathbf{\Gamma}_4+\mathbf{\Gamma}_5$	$\mathbf{\Gamma}_4+\mathbf{\Gamma}_5+2\mathbf{\Gamma}_6$
$C_{2\mathrm{v}}: E(z)$	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_1+\mathbf{\Gamma}_3$	$\mathbf{\Gamma}_2+\mathbf{\Gamma}_3+\mathbf{\Gamma}_4$	$\mathbf{\Gamma}_1+\mathbf{\Gamma}_2+\mathbf{\Gamma}_4$	$\mathbf{\Gamma}_5$	$\mathbf{\Gamma}_5$	$2\mathbf{\Gamma}_5$
$C_s: E(v): H(v)$	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_1+\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_1+2\mathbf{\Gamma}_2$	$2\mathbf{\Gamma}_1+\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_3+\mathbf{\Gamma}_4$	$\mathbf{\Gamma}_3+\mathbf{\Gamma}_4$	$2\mathbf{\Gamma}_3+2\mathbf{\Gamma}_4$

Table B.14 Compatibility table of the irreducible representations of the full rotation group for $L_z \leq 6$ and $J_z \leq \frac{13}{2}$ with the representations of the point group T_d

D_0^+	$\mathbf{\Gamma}_1$	D_0^-	$\mathbf{\Gamma}_2$
D_1^+	$\mathbf{\Gamma}_4$	D_1^-	$\mathbf{\Gamma}_5$
D_2^+	$\mathbf{\Gamma}_3+\mathbf{\Gamma}_5$	D_2^-	$\mathbf{\Gamma}_3+\mathbf{\Gamma}_4$
D_3^+	$\mathbf{\Gamma}_2+\mathbf{\Gamma}_4+\mathbf{\Gamma}_5$	D_3^-	$\mathbf{\Gamma}_1+\mathbf{\Gamma}_4+\mathbf{\Gamma}_5$
D_4^+	$\mathbf{\Gamma}_1+\mathbf{\Gamma}_3+\mathbf{\Gamma}_4+\mathbf{\Gamma}_5$	D_4^-	$\mathbf{\Gamma}_2+\mathbf{\Gamma}_3+\mathbf{\Gamma}_4+\mathbf{\Gamma}_5$
D_5^+	$\mathbf{\Gamma}_3+2\mathbf{\Gamma}_4+\mathbf{\Gamma}_5$	D_5^-	$\mathbf{\Gamma}_3+\mathbf{\Gamma}_4+2\mathbf{\Gamma}_5$
D_6^+	$\mathbf{\Gamma}_1+\mathbf{\Gamma}_2+\mathbf{\Gamma}_3+\mathbf{\Gamma}_4+2\mathbf{\Gamma}_5$	D_6^-	$\mathbf{\Gamma}_1+\mathbf{\Gamma}_2+\mathbf{\Gamma}_3+2\mathbf{\Gamma}_4+\mathbf{\Gamma}_5$
$D_{1/2}^+$	$\mathbf{\Gamma}_6$	$D_{1/2}^-$	$\mathbf{\Gamma}_7$
$D_{3/2}^+$	$\mathbf{\Gamma}_8$	$D_{3/2}^-$	$\mathbf{\Gamma}_8$
$D_{5/2}^+$	$\mathbf{\Gamma}_7+\mathbf{\Gamma}_8$	$D_{5/2}^-$	$\mathbf{\Gamma}_6+\mathbf{\Gamma}_8$
$D_{7/2}^+$	$\mathbf{\Gamma}_6+\mathbf{\Gamma}_7+\mathbf{\Gamma}_8$	$D_{7/2}^-$	$\mathbf{\Gamma}_6+\mathbf{\Gamma}_7+\mathbf{\Gamma}_8$
$D_{9/2}^+$	$\mathbf{\Gamma}_6+2\mathbf{\Gamma}_8$	$D_{9/2}^-$	$\mathbf{\Gamma}_7+2\mathbf{\Gamma}_8$
$D_{11/2}^+$	$\mathbf{\Gamma}_6+\mathbf{\Gamma}_7+2\mathbf{\Gamma}_8$	$D_{11/2}^-$	$\mathbf{\Gamma}_6+\mathbf{\Gamma}_7+2\mathbf{\Gamma}_8$
$D_{13/2}^+$	$\mathbf{\Gamma}_6+2\mathbf{\Gamma}_7+2\mathbf{\Gamma}_8$	$D_{13/2}^-$	$2\mathbf{\Gamma}_6+\mathbf{\Gamma}_7+2\mathbf{\Gamma}_8$

Table B.15 Character table and basis functions of the point group O_h (or m3m) which contains the inversion and consequently has parity as a good quantum number

O_h			E	$\bar{E}$	$8C_3$	$8\bar{C}_3$	$3C_2$ $3\bar{C}_2$	$6C_4$	$6\bar{C}_4$	$6C_{2'}$ $6\bar{C}_{2'}$	I	$\bar{I}$	$8S_6$	$8\bar{S}_6$	$3\sigma_h$ $3\bar{\sigma}_h$	$6S_4$	$6\bar{S}_4$	$6\sigma_d$ $6\bar{\sigma}_d$	Time reversal	Basis functions
A_{1g}	$\mathbf{\Gamma}_1$	$\mathbf{\Gamma}_1^+$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	a	R
A_{2g}	$\mathbf{\Gamma}_2$	$\mathbf{\Gamma}_2^+$	1	1	1	1	–1	–1	–1	1	1	1	1	1	1	–1	–1	–1	a	$(x^2-y^2)(y^2-z^2)(z^2-x^2)$
E_g	$\mathbf{\Gamma}_{12}$	$\mathbf{\Gamma}_3^+$	2	2	–1	–1	2	0	0	0	2	2	–1	–1	2	0	0	0	a	$(2z^2-x^2-y^2)$, $\sqrt{3}(x^2-y^2)$
F_{1g}	$\mathbf{\Gamma}'_{15}$	$\mathbf{\Gamma}_4^+$	3	3	0	0	–1	1	1	–1	3	3	0	0	–1	1	1	–1	a	S_x, S_y, S_z
F_{2g}	$\mathbf{\Gamma}'_{25}$	$\mathbf{\Gamma}_5^+$	3	3	0	0	–1	–1	–1	1	3	3	0	0	–1	–1	–1	1	a	yx, xz, xy
A_{1u}	$\mathbf{\Gamma}'_1$	$\mathbf{\Gamma}_1^-$	1	1	1	1	1	1	1	1	–1	–1	–1	–1	–1	–1	–1	–1	a	$\mathbf{\Gamma}_2^- \times \mathbf{\Gamma}_2^+$
A_{2u}	$\mathbf{\Gamma}'_2$	$\mathbf{\Gamma}_2^-$	1	1	1	1	1	–1	–1	–1	–1	–1	–1	–1	–1	1	1	1	a	xyz
E_u	$\mathbf{\Gamma}'_{12}$	$\mathbf{\Gamma}_3^-$	2	2	–1	–1	2	0	0	0	–2	–2	1	1	–2	0	0	0	a	$\mathbf{\Gamma}_3^+ \times \mathbf{\Gamma}_2^-$
F_{1u}	$\mathbf{\Gamma}_{15}$	$\mathbf{\Gamma}_4^-$	3	3	0	0	–1	1	1	–1	–3	–3	0	0	1	–1	–1	1	a	$\mathbf{\Gamma}_5^+ \times \mathbf{\Gamma}_1^-$
F_{2u}	$\mathbf{\Gamma}_{25}$	$\mathbf{\Gamma}_5^-$	3	3	0	0	–1	–1	–1	1	–3	–3	0	0	1	1	1	–1	a	$\mathbf{\Gamma}_5^+ \times \mathbf{\Gamma}_1^-$
		$\mathbf{\Gamma}_6^+$	2	–2	1	–1	0	$\sqrt{2}$	$-\sqrt{2}$	0	2	–2	1	–1	0	$\sqrt{2}$	$-\sqrt{2}$	0	a	$\phi(1/2,-1/2)$, $\phi(1/2,1/2)$
		$\mathbf{\Gamma}_7^+$	2	–2	1	–1	0	$-\sqrt{2}$	$\sqrt{2}$	0	2	–2	1	–1	0	$-\sqrt{2}$	$\sqrt{2}$	0	a	$\mathbf{\Gamma}_6^+ \times \mathbf{\Gamma}_2^+$
		$\mathbf{\Gamma}_8^+$	4	–4	–1	1	0	0	0	0	4	–4	–1	1	0	0	0	0	a	$\phi(3/2,-3/2)$, $\phi(3/2,3/2)$, $\phi(3/2,1/2)$, $\phi(3/2,3/2)$
		$\mathbf{\Gamma}_6^-$	2	–2	1	–1	0	$\sqrt{2}$	$-\sqrt{2}$	0	–2	2	–1	1	0	$-\sqrt{2}$	$\sqrt{2}$	0	c	$\mathbf{\Gamma}_6^+ \times \mathbf{\Gamma}_1^-$
		$\mathbf{\Gamma}_7^-$	2	–2	1	–1	0	$-\sqrt{2}$	$\sqrt{2}$	0	–2	2	–1	1	0	$\sqrt{2}$	$-\sqrt{2}$	0	c	$\mathbf{\Gamma}_6^+ \times \mathbf{\Gamma}_2^-$
		$\mathbf{\Gamma}_8^-$	4	–4	–1	1	0	0	0	0	–4	4	1	–1	0	0	0	0	c	$\mathbf{\Gamma}_8^+ \times \mathbf{\Gamma}_1^-$

Table B.16 The compatibility table of the irreducible representations of the group O_h and the irreducible representations of its proper subgroups

O^h	Γ_1^+	Γ_2^+	Γ_3^+	Γ_4^+	Γ_5^+	Γ_1^-	Γ_2^-	Γ_3^-	Γ_4^-
O	Γ_1	Γ_2	Γ_3	Γ_4	Γ_5	Γ_1	Γ_2	Γ_3	Γ_4
T_d	Γ_1	Γ_2	Γ_3	Γ_4	Γ_5	Γ_2	Γ_1	Γ_3	Γ_5
T_h	Γ_1^+	Γ_1^+	$\Gamma_2^+ + \Gamma_3^+$	Γ_4^+	Γ_4^+	Γ_1^-	Γ_1^-	$\Gamma_2^- + \Gamma_3^-$	Γ_4^-
D_{4h}	Γ_1^+	Γ_3^+	$\Gamma_1^+ + \Gamma_3^+$	$\Gamma_2^+\Gamma_5^+$	$\Gamma_4^+ + \Gamma_5^+$	Γ_1^-	Γ_3^-	$\Gamma_1^- + \Gamma_3^-$	$\Gamma_2^- + \Gamma_5^-$
D_{3d}	Γ_1^+	Γ_2^+	Γ_3^+	$\Gamma_2^+\Gamma_3^+$	$\Gamma_1^- + \Gamma_3^-$	Γ_1^-	Γ_2^-	Γ_3^-	$\Gamma_2^- + \Gamma_3^-$
$C_{4h} : H(z)$	Γ_1^+	Γ_2^+	$\Gamma_1^+ + \Gamma_2^+$	$\Gamma_1^+ + \Gamma_3^+ + \Gamma_4^+$	$\Gamma_2^+ + \Gamma_3^+ + \Gamma_4^+$	Γ_1^-	Γ_2^-	$\Gamma_1^- + \Gamma_2^-$	$\Gamma_1^- + \Gamma_3^- + \Gamma_4^-$
$C_{2h} : H(v)$	Γ_1^+	Γ_2^+	$\Gamma_1^+ + \Gamma_2^+$	$2\Gamma_2^+ + \Gamma_1^+$	$2\Gamma_1^+ + \Gamma_2^+$	Γ_1^-	Γ_2^-	$\Gamma_1^- + \Gamma_2^-$	$2\Gamma_2^- + \Gamma_1^-$
$C_{31} : H(w)$	Γ_1^+	Γ_1^+	$\Gamma_2^+ + \Gamma_3^+$	$\Gamma_1^+ + \Gamma_2^+ + \Gamma_3^+$	$\Gamma_1^+ + \Gamma_2^+ + \Gamma_3^+$	Γ_1^-	Γ_1^-	$\Gamma_2^- + \Gamma_3^-$	$\Gamma_1^- + \Gamma_2^- + \Gamma_3^-$
$C_{4v} : E(z)$	Γ_1	Γ_3	$\Gamma_1 + \Gamma_3$	$2\Gamma_2 + \Gamma_5$	$2\Gamma_4 + \Gamma_5$	Γ_2	Γ_4	$\Gamma_2 + \Gamma_4$	$\Gamma_1 + \Gamma_5$
$C_{2v} : E(v)$	Γ_1	Γ_2	$\Gamma_1 + \Gamma_2$	$2\Gamma_2 + \Gamma_3 + \Gamma_4$	$2\Gamma_1 + \Gamma_3 + \Gamma_4$	Γ_3	Γ_4	$\Gamma_3 + \Gamma_4$	$\Gamma_1 + \Gamma_2 + \Gamma_4$
$C_{3v} : E(w)$	Γ_1	Γ_2	Γ_3	$2\Gamma_2 + \Gamma_3$	$2\Gamma_1 + \Gamma_3$	Γ_2	Γ_1	Γ_3	$\Gamma_1 + \Gamma_3$

O^h	Γ_5^-	Γ_6^+	Γ_7^+	Γ_8^+	Γ_6^-	Γ_7^-	Γ_8^-
O	Γ_5	Γ_6	Γ_7	Γ_8	Γ_6	Γ_7	Γ_8
T_d	Γ_4	Γ_6	Γ_7	Γ_8	Γ_7	Γ_6	Γ_8
T_h	Γ_4^-	Γ_5^+	Γ_5^+	$\Gamma_6^+ + \Gamma_7^+$	Γ_5^-	Γ_5^-	$\Gamma_6^- + \Gamma_7^-$
D_{4h}	$\Gamma_4^- + \Gamma_5^-$	Γ_6^+	Γ_7^+	$\Gamma_6^+ + \Gamma_7^+$	Γ_6^-	Γ_7^-	$\Gamma_6^- + \Gamma_7^-$
D_{3d}	$\Gamma_1^- + \Gamma_3^-$	Γ_4^+	Γ_4^-	$\Gamma_6^+ + \Gamma_5^- + \Gamma_6^-$	Γ_4^-	Γ_4^-	$\Gamma_4^- + \Gamma_5^- + \Gamma_6^-$
$C_{4h} : H(z)$	$\Gamma_2^- + \Gamma_3^- + \Gamma_4^-$	$\Gamma_5^+ + \Gamma_6^+$	$\Gamma_7^+ + \Gamma_8^+$	$\Gamma_5^+ + \Gamma_6^+ + \Gamma_7^+ + \Gamma_8^+$	$\Gamma_5^- + \Gamma_6^-$	$\Gamma_8^- + \Gamma_7^-$	$\Gamma_5^- + \Gamma_6^- + \Gamma_8^- + \Gamma_7^-$
$C_{2h} : H(v)$	$2\Gamma_1^- + \Gamma_2^-$	$\Gamma_3^+ + \Gamma_4^+$	$\Gamma_3^+ + \Gamma_4^+$	$2\Gamma_3^+ + 2\Gamma_4^+$	$\Gamma_3^- + \Gamma_4^-$	$\Gamma_3^- + \Gamma_4^-$	$2\Gamma_3^- + 2\Gamma_4^-$
$C_{3l} : H(w)$	$\Gamma_1^- + \Gamma_2^- + \Gamma_3^-$	$\Gamma_4^- + \Gamma_5^-$	$\Gamma_4^- + \Gamma_5^-$	$\Gamma_4^+ + \Gamma_5^+ + 2\Gamma_6^+$	$\Gamma_4^- + \Gamma_5^-$	$\Gamma_4^- + \Gamma_5^-$	$2\Gamma_4^+ + \Gamma_5^+ + 2\Gamma_6^+$
$C_{4v} : E(z)$	$\Gamma_3 + \Gamma_5$	Γ_6	Γ_7	$\Gamma_6 + \Gamma_7$	Γ_6	Γ_7	$\Gamma_6 + \Gamma_7$
$C_{2v} : E(v)$	$\Gamma_1 + \Gamma_2 + \Gamma_3$	Γ_5	Γ_5	$2\Gamma_5$	Γ_5	Γ_5	$2\Gamma_5$
$C_{3v} : E(w)$	$\Gamma_2 + \Gamma_3$	Γ_4	Γ_4	$\Gamma_4^+ + \Gamma_5^+ + 2\Gamma_6^+$	Γ_4	Γ_4^+	$\Gamma_4 + \Gamma_5 + \Gamma_6$

Problems

B.1 Calculate the symmetries of the states which can be reached by one- and two-photon transitions in point groups O_h and T_d.

B.2 Calculate the possible symmetries of the excitons with principal quantum number $n_\mathrm{B} = 1, 2$ and 3 in the point groups O_h, T_d and $C_{6\mathrm{v}}$.

B.3 Calculate the symmetries of the ground state ($\mathbf{\Gamma}_1$ envelope) of $\mathrm{D}^0\mathrm{X}$, $\mathrm{A}^0\mathrm{X}$ and $\mathrm{D}^+\mathrm{X}$ centers in O_h, T_d, and $C_{6\mathrm{v}}$, assuming that the radius of these complexes is so large that the carriers feel the full symmetry of the lattice and not only the nearest neighbors.

B.4 SnO_2 is a direct gap semiconductor with the band extrema at the $\mathbf{\Gamma}$-point. It crystallizes in the point group D_{4h}. Which symmetries do you expect for the highest valence and the lowest conduction band. The lowest band-to-band transition is dipole forbidden. Why? Do you except dipole-allowed exciton transitions? Do you think that they will have a big oscillator strength?

B.5 Rock salt, diamond and Cu_2O all have point group O_h (i.e., a face-centered-cubic lattice), but different crystal structures. Inspect the crystal structures. Find the primitive unit cell and the nonprimitive cubic one. Parity is a good quantum number. The highest valence band is formed in Cu_2O from the $3d$ levels of Cu^+ and the lowest conduction bands result from the Cu^+ $4s$ and $4p$ levels. Which symmetries do you expect at the $\mathbf{\Gamma}$-point and which exciton states with S envelope? Are they dipole allowed or forbidden?

References

[60H1] V. Heine, *Group Theory in Quantum Mechanics* (Pergamon, Oxford, 1960)

[63K1] G.F. Koster et al. (eds.), *Properties of the Thirty-Two Point Groups* (MIT, Cambridge, 1963)

[64K1] R.S. Knox, A. Gold, *Symmetry in the Solid State* (Benjamin, New York, 1964)

[64T1] W.A. Tinkham, *Group Theory in Quantum Mechanics* (McGraw-Hill, New York, 1964)

[67S1] H.W. Streitwolf, *Gruppentheorie in der Festkörperphysik* (Akademie-Verlag, Leipzig, 1967)

[74B1] G.L. Bir, G.E. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors* (Wiley, New York, 1974)

[76C1] K. Cho, Phys. Rev. B **14**, 4463 (1976)

[77G1] O. Goede, Phys. Status Solidi (b) **81**, 235 (1977)

[85H1] B. Hönerlage et al., Phys. Rep. **124**, 161 (1985)

[88L1] W. Ludwig, C. Falter, *Symmetries in Physics*, Springer Series in Solid-State Science, vol. 64 (Springer, Berlin, 1988)

[90I1] T. Inui, Y. Tanabe, Y. Onadera, *Group Theory and Its Applications in Physics*. Springer Series on Solid State Sciences, vol. 78 (Springer, Berlin, 1990)

[92I1] T. Hahn, *International Tables for Crystallography Vol. A Space Groups*, 3rd edn. (Kluwer, Dordrecht, 1992)

[93J1] S. Jorda, U. Rössler, D. Broido, Phys. Rev. B **48**, 1669 (1993)

[94H1] Th. Hahn, H. Wondratschek, *Symmetry of Crystals*. Lecture Notes to the Summer School in Gjuletchitsa, Bulgaria (Heron Press Ltd., Sophia, 1994)

[96I1] E.L. Ivchenko, Yu.A. Kaminski, U. Rössler, Phys. Rev. B **54**, R2315 (1996)

[97I1] E.L. Ivchenko, G.E. Pikus, *Superlattices and Other Heterostructures*, 2nd edn. (Springer, Berlin, 1997)

[00S1] M. Schmidt et al., Appl. Phys. Lett. **77**, 85 (2000)

[01L1] U. Rössler (ed.), *Landolt-Börnstein*. New Series, Group III, vol. 41a1, part 2 (Springer, Berlin, 2001)

[02M1] A.A. Maksimov et al., Phys. Status Solidi **229**, 35 (2002)

[03T1] E. Tsitisishvili, *Spectroscopy of Systems with Spatially Confined Structures*. NATO Science Series II, ed. by B. Di Bartolo, vol. 90 (Kluwer, Dordrecht, 2003), p. 357

[10Y1] P.Y. Yu, M. Cardona, *Fundamentals of Semiconductors*, 4th edn. (Springer, Berlin, 2010)

[18H1] B. Hönerlage, I. Pelant, *Symmetry and Symmetry-Breaking in Semiconductors: Fine Structure of Exciton States*. Springer Tracts in Modern Physics, vol. 279 (Springer International Publishing AG 2018)

Appendix C
Important Data of Semiconductors

Abstract In the tables of this appendix (Tables C.1 and C.2) we summarize some band parameters of semiconductors. An exhaustive listing of data is found in [82L1].

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0

Table C.1 Values of the band gap energy of some group IV, III–V, II–VI and I–VII semiconductors (SC) at low temperature ($T \lesssim 10$ K), Sy: Symmetry (point group) (According to [82L1] and [01V1, 04A1])

Group IV				Group III–V				Group II–VI				Group I–VII			
SC	Sy	E_g (eV)	dir/indir	SC	Sy	E_g (eV)	dir/indir	SC	Sy	E_g (eV)	dir/indir	SC	Sy	E_g (eV)	dir/indir
C	O_h	5.48	i	AlN	C_{6v}	6.28	d	ZnO	C_{6v}	3.437	d	CuCl	T_d	3.395	d
S*i*	O_h	1.17	i	AlP	T_d	2.53	i	ZnS	C_{6v}	3.91	d	CuBr	T_d	3.077	d
Ge	O_h	0.744	i	AlAs	T_d	2.228	i	ZnS	T_d	3.78	d	CuI	T_d	3.115	d
Grey Sn	O_h	0	Semimetal	AlSb	T_d	1.696	i	ZnSe	T_d	2.82	d	AgCl	T_d	3.249	i
				GaN	C_{6v}	3.503	d	ZnTe	T_d	2.391	d	AgBr	T_d	2.684	i
				GaP	T_d	2.350	i	CdO	O_h	0.8	i	AgI	C_{6v}	3.024	d
				GaAs	T_d	1.518	d	CdS	C_{6v}	2.583	d				
				GaSb	T_d	0.812	d	CdSe	C_{6v}	1.841	d				
				InN	C_{6v}	0.67[a]	d	CdTe	T_d	1.60	d				
				InP	T_d	1.424	d	HgS	T_d	0	d				
				InAs	T_d	0.418	d	HgSe	T_d	0	d				
				InSb	T_d	0.237	d	HgTe	T_d	0	d				

[a]New data, see [02D1, 02K1]

Table C.2 Effective masses of some selected semiconductors (From [82L1])

SC	Sy	dir/ind	m_e/m_0	m_{hh}/m_0	m_{lh}/m_0	$m_A \perp \parallel /m_0$
C	T_d	i		2.18	0.7	
Si	T_d	i	$\perp 0.19$ $\parallel 0.92$	0.54	0.15	
Ge	T_d	i	$\perp 0.081$ $\parallel 1.6$	0.3	0.043	
AlAs	T_d	i	$\perp 1.56, 5.8$ $\parallel 0.19$	0.76	0.15	
AlSb	T_d	i	$\perp 0.26$ $\parallel 1.0$	0.94	0.11	
GaN	C_{6v}	d	0.22			≈ 0.8
GaP	T_d	i	$\perp 0.25$ $\parallel 7.25; 2.2$	0.6	0.17	
GaAs	T_d	d	0.066	0.47	0.07	
GaSb	T_d	d	0.042	0.35	0.05	
ZnO	C_{6v}	d	0.28			$\perp 0.45$ $\parallel 0.59$
ZnS	C_{6v}	d	0.28			0.5
ZnSe	T_d	d	0.15	0.8	0.145	
ZnTe	T_d	d	0.12	0.6		
CdS	C_{6v}	d	0.2			$\perp 0.7$ $\parallel 2.5$
CdSe	C_{6v}	d	0.13			$\perp 0.45$ $\parallel 1.1$
CdTe	T_d	d	0.1	0.4		
CuCl	T_d	d	0.4	2.4		
CuBr	T_d	d	0.25	1.4		
CuI	T_d	d	0.3	≈ 2		

References

[82L1] O. Madelung, U. Rössler (ed.), *Landolt–Börnstein*. New Series, Group III, vol. 17a to i, 22a and b, 41A to D (Springer, Berlin, 1982–2001)

[01V1] I. Vurgaftman, J.R. Meyer, L.R. Ram-Mohan, J. Appl. Phys. **89**, 5815 (2001)

[02D1] V.Yu. Davidov et al., Phys. Status Solidi B **234**, 787 (2002)

[02K1] A. Klochikhin et al., Phys. Status Solidi b **234** 787 (2002); Phys. Rev. B **71** 195207 (2005)

[04A1] B. Arnaudov et al., Phys. Rev. B **69**, 115216 (2004)

Appendix D
Growth and Structural Properties of Low-Dimensional Semiconductor Structures

Abstract Here we want to give a short review of the most widely used techniques for the epitaxy of low-dimensional semiconductor structures like quantum wells (QW), superlattices (SL), quantum wires (QWR) and quantum dots (QD). For more detailed information the reader is referred to, e.g., [80L1, 90S1, 01L1, 93C1, 00K2, 03T1, 07N1, 10G1, 13L1]. For graphene or other monolayer semiconductors, see, e.g., [08S1, 09N1, 14D1, 16N1]. We also discuss some important structural properties of common semiconductors.

D.1 Quantum Wells and Superlattices

The two most important ways to grow thin (quantum)films and related (nano)-structures are *molecular beam epitaxy (MBE)* and *metal-organic chemical vapour deposition (MOCVD)* also known as *metal-organic vapour phase epitaxy (MOVPE)* (see Fig. D.1). Further techniques are hot wall (beam) epitaxy (HW(B)E) or pulsed laser deposition (PLD). For a few reference for these techniques see [92B1, 93G1, 03T1, 08L1, 10L1, 13L1].

In MBE one creates, in an ultra-high vacuum chamber, atomic or molecular beams of the constituents of the semiconductor layers to be grown from heated effusion cells controlled by mechanical shutters. By simultaneously or alternatively opening the shutters of, e.g., Ga and As cells it is possible to grow a GaAs layer on a suitable, heated, mono-crystalline substrate and with the Al and As cells an AlAs layer, respectively. Doping atoms can be incorporated opening further cells, e.g., Si or Be for n and p doping of III–V compounds. In II–VI epitaxy one sometimes also uses compound sources, e.g., for CdS. Some modifications of the MBE process are known as *atomic layer epitaxy (ALE)* or *migration enhanced epitaxy (MEE)*, which allow one to grow layers controlled on a monolayer scale.

In MBE the atomic or molecular beams propagate collision-free from the effusion cells through the ultrahigh vacuum to the substrate where they are physisorbed and

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0

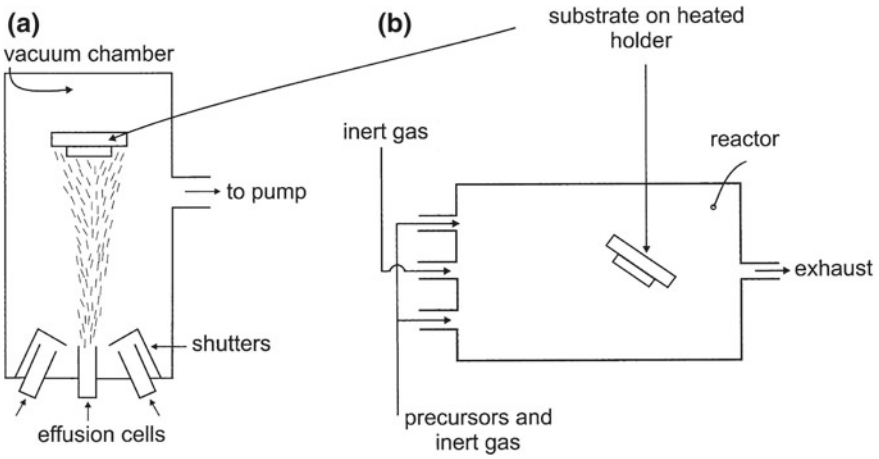


Fig. D.1 Schematics of an MBE chamber (**a**) and of an MOCVD reactor (**b**)

then either chemisorbed and integrated in the growing surface or desorbed again. The MBE technique operates far from thermodynamic equilibrium.

Growth by HWE/HWBE occurs, in contrast, close to equilibrium. In this technique the substrate is mounted directly on top of the effusion cell or very close to it. The source is kept at a temperature T_{so}, the walls of the growth chamber at higher temperature $T_w > T_{so}$ to avoid condensation of the source material at the walls and the substrate is kept at slightly lower temperature $T_{sub} < T_{so}$, to allow a net flow of material from the source to the substrate.

In MOCVD so-called *precursors* like AsH_3 and $Ga(CH_3)_3$ are brought with an inert gas like N_2 as a transport gas into the reactor. A chemical reaction ideally takes place directly on the heated substrate. In this example, e.g.,

$$AsH_3 + Ga(CH_3)_3 \rightarrow GaAs + 3CH_4 \tag{D.1}$$

resulting in the deposition of GaAs.

Important for epitaxial growth of two materials on top of each other with good quality are the following criteria. Ideally the materials should should have:

- equal or similar crystal structures, e.g., both should crystallize in the zincblende structure or should show at least the same coordination like a tetrahedral one in diamond, zincblende and wurtzite structure.
- almost identical lattice constants. Below we discuss consequences resulting from a lattice mismatch,
- the same valences, which means that III–V semiconductors are more easy to grow on a III–V substrate than on II–VI or group IV substrates and vice versa,
- similar types of chemical binding, e.g., either preferentially covalent or ionic.

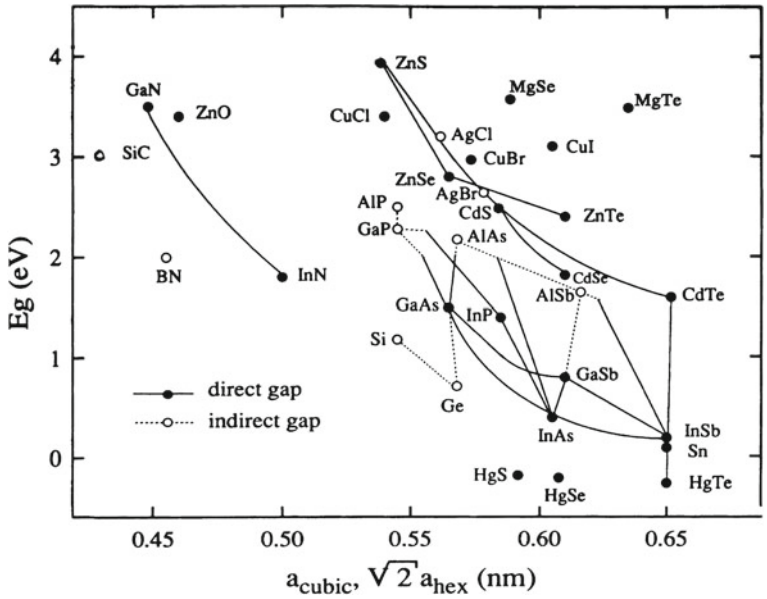


Fig. D.2 Width of the band gaps as a function of the lattice constant for various semiconductors and their alloys. *Solid lines*: direct gap; *dashed lines*: indirect gap. Note that this figure gives information on the band gap only, but not on the relative band alignments, i.e., whether a combination will be of type I, II or III (Data from [82L1, 01K1])

The trivial way to fulfill these conditions is *homoepitaxy*, i.e., to grow on a substrate the same material. This partly allows producing high quality, e.g., chemically very pure layers on a less perfect substrate of the same material. As one out of the many examples see [07N1]. However, it was found e.g., in ZnO that even homoepitaxy may cause some problems. See Chap. 3 of [10K2].

In contrast, substrate and growing layers are different in *heteroepitaxy*. As can be seen in Fig. D.2 the system GaAs, AlAs and their alloys $Al_xGa_{1-x}As$, which exist for all values $0 \leq x \leq 1$, fulfill these conditions in an almost ideal manner. The band alignment is of type I where GaAs forms the well with about 60% of the band discontinuity in the conduction band. GaAs is a direct-gap semiconductor, AlAs an indirect one. The transition occurs in the alloy for $x \approx 0.43$ (see [96K1, 01V1]). To avoid complications by combining a direct-gap material with an indirect one on one side and to have a substantial confinement potential on the other, one usually limits x to values $0.3 \leq x \leq 0.4$.

Since GaAs substrates are available in the highest quality compared to all other substrate materials, one tries to use it and to choose the compositions of the layers in a way to match the lattice constants like in $In_{0.51}Al_{0.49}P$, $AlAs_{0.96}P_{0.04}$ or $Ga_{0.51}In_{0.49}P$. InP is another widely used substrate to which III–V alloys can be lattice matched like $Al_{0.48}In_{0.52}As$, $GaAs_{0.5}Sb_{0.5}$ or $Ga_{0.47}In_{0.53}As$. Quaternary alloys

like $In_{1-y}Ga_yAs_{1-x}P_x$ allow varying the lattice constant and the band gap to some extent independently as can be seen from the border lines Fig. D.2.

A further III–V system, which has become rather popular in recent years due to the success of building blue- and UV-light emitting and laser diodes, is GaN including its alloys with InN and AlN. Due to the huge lattice mismatch to GaAs, this system is grown on SiC, ZnO or most commonly on Al_2O_3.

The II–VI compound ZnSe is in most cases grown on GaAs due to the almost perfect lattice match and in spite of the problems resulting from the interface between a III–V and a II–VI compound [00N1]. ZnSe-based quantum structures including $Zn_{1-y}Cd_ySe$ or $ZnSe_{1-x}Te_x$ wells and $Zn_{1-y}Mg_ySe_{1-x}S_x$ barriers are type-I structures whose luminescence covers the blue-green to yellow spectral region.

Another quantum-well system for the blue/near UV is based on ZnO including its alloys with MgO, BeO and CdO [07K1, 09W1, 10A1, 10K1, 10K2, 10L1, 13L1].

In many cases it is unavoidable, but sometimes even intentional, to grow materials on each other, that have different lattice constants. Prominent examples are $In_{1-y}Ga_yAs$ on GaAs, $ZnSe_{1-x}Te_x$ wells on ZnSe and type II CdS/ZnSe or CdSe/CdS superlattices [01L2]. Another important *strained-layer superlattice* system is formed by the group V elements C, Si and Ge [07L1, 13L1].

Such a lattice mismatch has consequences for the growing layer. If the growing layer wets the substrate, it will start to grow as a tensile or compressively strained layer, depending on the sign of the mismatch. In addition, tensile in-plane strain usually results in a uniaxial compressive distortion of the lattice in growth direction and vice versa.

The elastic energy stored in the strained layer increases with layer thickness until the lattice relaxes at a *critical thickness* l_c to its own lattice constant. This relaxation can result in the formation of *dislocations* like in CdS on GaAs [94G1, 97H1]. The formation of dislocations is highly undesirable because they form scattering and non-radiative recombination centers, degrading the in-plane carrier mobility and the luminescence efficiency. The formation (and multiplication) of dislocations in strained ZnSe-based laser diode structures limits their lifetime to values unacceptable for commercial devices.

Another possibility is that the growing layer starts to form on the wetting layer small islands, in which it can relax toward its own lattice constant. This growth mode is known as *Stranski-Krastanov growth mode* [99B1] and is a standard route to realize *self-assembled quantum dots (SAQD)*. However, these islands are subject to segregation and diffusion during the necessary overgrowth with barrier material [10G1]. In contrast the ideal two-dimensional layer by layer growth is known as *Frank-Van der Merwe growth mode*, and the direct formation of island without wetting layer is called *Vollmer-Weber growth mode* [10G1].

As a consequence, strained layers are usually grown only below the critical thickness l_c in (M)QW or SL. Superlattices with many periods (e.g., 50–100) may adapt, after a few periods, an intermediate lattice constant which results in two-dimensional compressive and tensile strain in the two different materials, respectively, with a complementary uniaxial distortion normal to the layers. These are so-called *free-standing superlattices*. An example are Ge/Si superlattices [07L1].

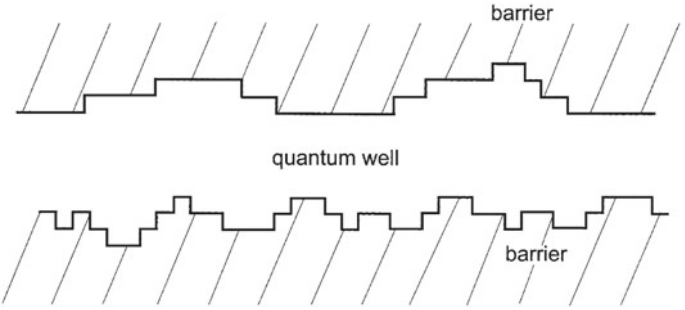


Fig. D.3 Thickness fluctuations of a quantum well with two different length scales on both sides (schematic)

Such strain has of course consequences for the electronic band structure as is illustrated in Fig. 14.9. For compressive strain the band gap widens and the hh valence band is closer to the conduction band. For tensile strain the situation the opposite. Depending on the sign of the strain this splitting may either enhance or partly compensate the splitting induced by quantization.

To conclude our discussion on layered structures we briefly mention some complications encountered in real quantum wells and superlattices. In spite of all progress in epitaxy, real quantum wells typically exhibit, at both interfaces fluctuations of at least one monolayer height as shown schematically in Fig. D.3. The length scale of the fluctuations may be different, e.g., compared to the excitonic Bohr radius and it may even depend on the growth sequence, i.e., if GaAs is grown on AlAs or vice versa.

An extreme case of interface fluctuations is give in *corrugated superlattices*. On highly indexed planes, like [311] in the AlAs/GaAs system, corrugated layers can be grown that have periodically modulated thickness. Actually it has been found that one interface is intermixed and the other is corrugated. For more details of these structures, which were also known as quantum well wire superlattices (QWWSL) see, e.g., [96K1, 98L1, 01K1].

The next problem is that the interfaces are usually not sharp but washed out by diffusion and segregation. These effects are, e.g., well known for InAs wells in GaAs or for CdSe wells in ZnSe (see, e.g., [97W1, 01K1, 13L1]. They can be influenced to some extent by the growth parameters.

In the caption of Fig. D.2 we noted that it gives the band gap of the various compounds, but not the band alignment. The band alignment depends trivially on the two materials but is often not very precisely known. In the case of alloys in wells or barriers the offset ratio of the conduction and valence band may depend on the composition of the alloy. It depends further on strain, on the crystallographic orientation of the two materials, on well and barrier thicknesses, on the termination of the interfaces especially in case of quaternary systems AB/CD without common anion or cation, on the presence of dipole layers at the interface, on intermixing

(see above) and even on the growth sequence, i.e., if material I is grown on top of material II or vice versa, so-called non-commutativity. For data on the offsets in III–V quantum-well structures see [01K1], for II–VI structures see [04K1].

It should be emphasized that all these types of disorder—like well-width fluctuations, intermixing or alloy formations—result in the formation of *localized states*, as explained in more detail in Sect. 18.2.

D.2 Monolayer Semiconductors

The fabrication of *monolayer semiconductors* seem on first sight simple and straight forward. But to achieve high quality layers with large lateral size, defined number of layers and without chemical modification (like oxidation, which can be beneficial for some applications) is a non- trivial task. Many procedures rely on the fact, that the binding is weak between the individual 2D layers. So, rubbing of a respective bulk material (like a graphite pencil) on a surface or repeated peeling with adhesive tape is used to produce layers of graphene or other 2D materials [07G1, 12W1]. This simple *micro-mechanical cleavage* typically leads to flakes of various thicknesses so that monolayers have to be identified. This is can be achieved by using an optical microscope since interference makes a monolayer material placed on an oxidized Si substrate visible [05N1]. Raman scattering is a method to identify the exact number of monolayers [13T1].

Since micro-mechanical cleavage is not suitable for mass production there is the search for alternative methods. Ultra-sonication in solvents called *liquid-phase exfoliation* followed by sedimentation-based ultra-centrifugation can lead to isolation of flakes with controlled thickness. Exfoliation of monolayers is very efficient after intercalation of TMDs with ions like Lithium in a solvent containing these ions. Subsequent exposure to water leads to a strong reaction with the ions separating the TMD layers. A drawback of these methods is that the layers are structurally and chemically modified [12W1]

Chemical vapor deposition on metal foils or SiC is used for graphene. It mainly yields multi-layers which can be grown on wafer scale [10B1, 12N1, 12W1]. Meanwhile sequential CVD of combinations of graphene and various monolayer semiconductors, direct CVD of TMD structures by vapor-solid reactions and layered growth by van der Waals epitaxy have been established (see [14D1, 16N1] for reviews).

D.3 Quantum Wires and Nanorods

The first attempts to create quantum wires were performed by lithography and etching of quantum wells. To reduce the density of surface states and thus detrimental non-radiative recombination the structures have to be overgrown by some barrier material.

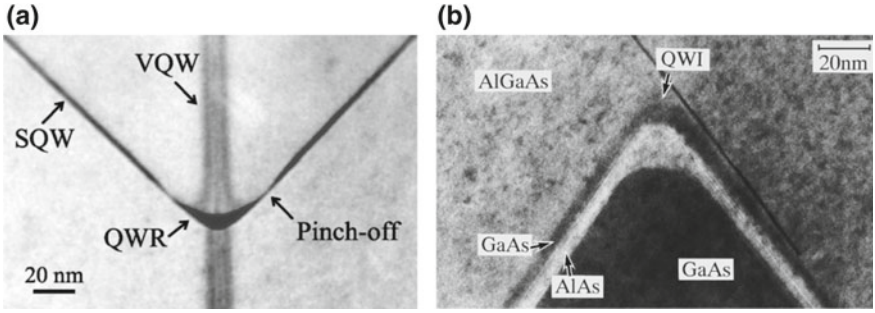


Fig. D.4 TEM image of a V-groove quantum wire. Overgrowth of a single quantum well (SQW) on a pre-patterned substrate leads to a crescent-shape quantum wire (QWR) and a vertical quantum well (VQW) [04C1] (**a**); TEM image of a GaAs ridge quantum wire (QWI) (**b**) (reprinted with permission from [96I1] ©1996 AIP Publishing https://doi.org/10.1063/1.360941)

By means of lithography and selective etching it is possible to create pre-patterned substrates with either V grooves (Fig. D.4a) or ridges (Fig. D.4b). A quantum well is now grown on these pre-patterned substrates, generally by MOCVD on the V-grooves and by MBE on the ridges. The enlarged part of this quantum well on the bottom of the V-groove or on the top surface of the ridge forms the confining potential in two dimensions.

In the case of V-grooves it is possible to stack several identical quantum wires on top of each other leading to *quantum-wire superlattices*. For more details see, e.g., [99B2, 00B1, 02K1, 13L1].

A further method to produce wires has been shown schematically in Fig. 17.14. Growth starts with the deposition of a quantum well. The sample is then cleaved in ultra high vacuum normal to this quantum well and another quantum well is grown on this cleaved surface followed by barrier material. This technique is known as *cleaved-edge overgrowth* [85C1, 96S1, 13L1]. In the junction region of the two quantum wells provides a confining potential in two directions. If one starts from a multiple quantum well or a superlattice one obtains by the cleaved edge overgrowth an array of uncoupled or coupled quantum wires, respectively.

High crystalline quality of *nano-wires* or *nanorods* can be achieved by self-organized growth using vapor-phase transport assisted by a catalyst [01H1]. This growth mechanism works for various semiconductor materials [09Y1] including Si [09S1]. We show here examples for ZnO in Fig. D.5. A thin layer of Au or a nano-patterned array of Au dots is first deposited on a substrate. Ordered arrays can be achieved by using various template techniques for the patterning (see, e.g., [06F1, 07Z1]). Depending on specifics like the choice of substrate and growth parameters growth of ZnO nano-wires or nanorods occurs at the position of the Au dots. The ZnO structures can be defect poor single crystals aligned in arrays. Sometimes Au dots are found on top of the ZnO rods indicating that growth occurred in the vapor-liquid-solid mechanism [64W1, 07R1]. In other cases ZnO nuclei form at the position of the Au dots from where the nanorod growth continues [06F1, 07Z1].

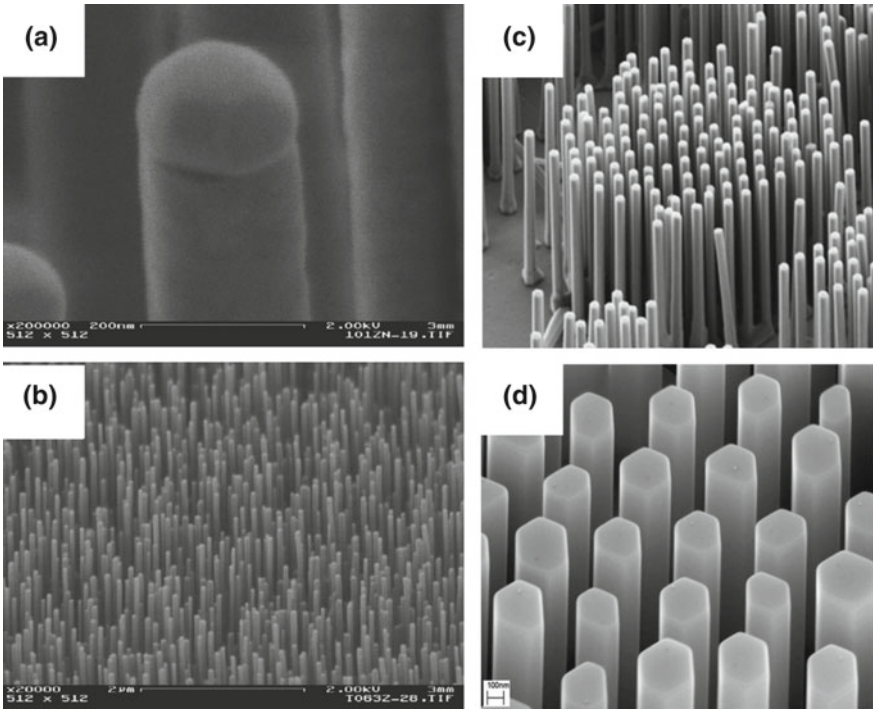


Fig. D.5 SEM image of a ZnO nanorod grown by VLS with the Au nano-dot on top (**a**); array of VLS-grown ZnO nanorods (**b**); ZnO nanorod array fabricated by catalytic growth on an ordered Au nano-dot array (**c**); hexagonal ZnO nanorods (**d**); (**a**, **b** reprinted with permission from [07R1] ©2007 AIP Publishing https://doi.org/10.1063/1.2710295, **c** reprinted with permission from [07Z1] ©2007 AIP Publishing https://doi.org/10.1063/1.2805073, **d** reprinted with permission from [08Z1] ©2008 AIP Publishing https://doi.org/10.1063/1.2907197)

Further, nano-wires from many different semiconductor materials can be grown using (seeded) solution-based methods (see, e.g., [02K2, 03M1, 08K1]). Here, core/shell structures can be realized and also higher order morphologies, i.e., branched structures like tetrapods occur.

Apart from nanorods frequently other nano-structures grow like nano-walls, belts, nails, tubes, flowers, dandelions, castles, corals or cabbage depending on the fantasy of the respective authors [07K1, 10K1, 10K2]. Frequently these nano-structures have limited optical quality or reproducibility [04L1].

For a review on use of nano-wires in photonic applications see [09Y1].

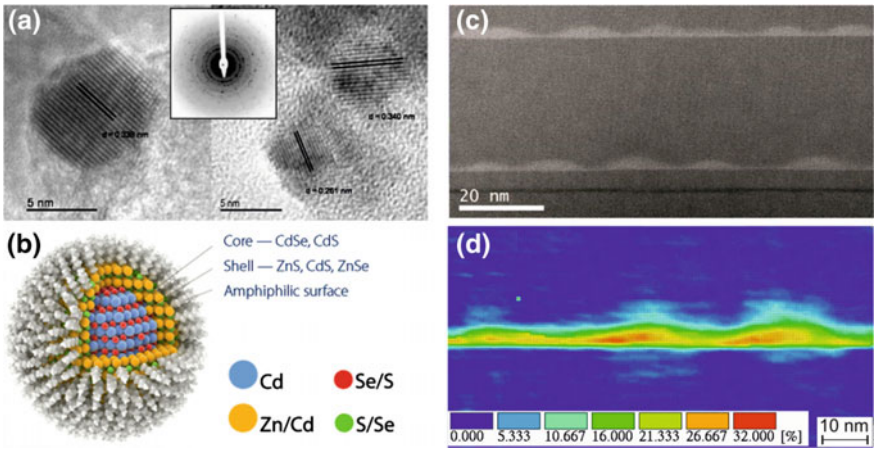


Fig. D.6 Various possibilities to form quantum dots: HRTEM micro-graph of spherical ZnS particles embedded in amorphous silica after annealing of the doped glasses. The inset shows the selected area electron diffraction (SAED) pattern reflecting the random orientation of the crystalline dots (republished with permission from [08P1] ©2008 Royal Society of Chemistry; permission conveyed through Copyright Clearance Center, Inc.) (**a**); schematic depiction of the structure of a II/VI core/shell quantum dot enclosed by organic ligands (http://en.rusnano.com/press-centre/news/88604 [19R1]) (**b**); HR-STEM micro-graph of two stacked layers of InAs/InP SAQDs (http://www.fotonik.dtu.dk/english/research/nanophotonics/nanodev/research/quantum-dots [19D1])(**c**); color-coded In-content of InGaAs/GaAs SAQDs deduced with CELFA from a HRTEM micrograph [18C1] (**d**)

D.4 Quantum Dots

Quantum dots (QD) can have a variety of manifestations derived from principally different fabrication routes [13L1]. Most important are colloidal QDs forming by precipitation, self-assembled quantum dots (SAQD) established in strained-layer overgrowth, QDs grown on pre-structured substrates, quasi-0D confining potentials (interfacial quantum dots (IQD)) realized by disorder, and electrostatically defined QDs. We will review here the respective fabrication techniques, except for the last case.

A technique to form QDs, which was established for a long time before recognizing the quasi-0D character of the electronic states, is an annealing procedure in glasses doped with the constituents of a semiconducting compound [85E1, 95W1]. This procedure is widely used for II–VI compounds such as $CdS_{1-x}Se_x$ including the cases $x = 0$ and $x = 1$, CdTe, ZnS or for the Cu halides. The commercially available glass edge filters are made in this way. Early examples are described in [59W1, 83Z1]. The average diameter of the dots increases with annealing time and temperature. We show in Fig. D.6 a high-resolution transmission electron (HRTEM) micro-graph of dots in a silicate glass matrix prepared by a sol-gel method. The lattice planes within the almost spherical nano-crystallite are nicely visible [08P1].

Colloidal QDs are produced by pyrolysis of organometallic precursors in hot (organic) liquids or gels [84B1, 93M1, 96A1, 08O1, 10H1, 10S1]. These techniques result in dots that are almost spherical and have a high barrier. The unavoidable size distribution of the dots can be reduced to a few percent by fractional sedimentation.

The nano-particles can be fabricated in the form of core/shell quantum dots like CdS/ZnS QDs (see Fig. D.6) to suppress blinking of the luminescence due to surface effects [13C1]. It is even possible to grow onion-like multiple quantum wells or superlattices around a core [13L1]. The nano-particles are typically covered by organic ligands to prevent agglomeration, to allow bio-safe application and/or to enable specific binding of the nano-particle to proteins by key-lock processes in marker applications [13K1].

Still another possibility is to grow the dots in the voids of zeolithe crystals or of synthetic opals [99G1].

The second important class of quasi-0D systems are *self-assembled quantum dots (SAQD)*. In strongly lattice mismatched systems (strong means a lattice misfit beyond about 5%), apart from the formation of misfit dislocations, there is the possibility of *Stranski-Krastanov growth* where small hillocks form on a wetting layer. In theses hillocks or small pyramids the growing compound relaxes towards its own lattice constant. The formation of these pyramids can be verified, e.g., by scanning force microscopy [99B1].

Due to the large surface area the pyramids are actually optically dead. The have to be capped again by the respective high-gap material to form quasi-0D structures which can be employed in optical devices. The preferential systems in which these self-assembled quantum dots are investigated are $In_{1-y}Ga_yAs/GaAs$ [99B1] and CdSe/ZnSe [04S1].

The SAQDs are unfortunately not perfect in composition and shape. It has been found by HRTEM investigations in both systems that frequently during overgrowth a substantial intermixing between well and barrier material occurs [99B1, 10G1]. The consequence is that the quantum islands consist in most cases of pronounced thickness or composition fluctuations rather than of well-defined pyramids. We give two examples in Fig. D.6c,d. Analysis of the HRTEM micro-graphs by methods like *Composition Evaluation by Lattice Fringe Analysis (CELFA)* revealed that typically, the SAQDs do not have the expected chemical composition [03R1]. But, the contrast in composition is still high and allows efficient quasi-0D carrier confinement in the local potential minima.

Even more complex are SAQDs from II–VI compounds. On ZnSe surfaces pyramids can form, e.g., in air even in the absence of CdSe [97H1, 97R1]. In the CdSe/CdS system scanning microscopy revealed also islands, which are most likely some Se precipitates [98G1]. Still it is possible to create quantum islands, which contain, at least in their center, CdZnSe with a high Cd-content [00K1].

It is also possible to stack these quantum islands with the dots aligned on top of each other. This is induced when the barrier layer in multiple quantum well (or island) systems is thin enough that the $(n + 1)$th layer feels the lattice distortion of the islands in the nth layer below. For examples of the InAs/GaAs and the CdSe/ZnSe systems, see, e.g., [99B1, 00K1, 01S1].

The random in-plane distribution of self-assembled quantum dots is hampering the individual addressability of the dots which is, e.g., desirable in quantum devices. Thus there are efforts to grow QDs on pre-structured substrates at well-defined positions. Corner-cube or inverted-pyramid holes or pyramids are used as well as quadratic hole arrays [97H2, 02K1, 11H1].

Another possibility to define local potential is the usage of so-called stressors. These are small areas of a non-lattice matched material grown on top of a thin barrier layer over a quantum well. The strain exerted by the stressors on the quantum well material must be such that the band gap in the well below the stressor is reduced creating a confining lateral potential [97W2].

The AlAs/GaAs system is not prone to significant inter-diffusion and segregation. But, thickness fluctuations occur readily in GaAs quantum wells as schematically shown in Fig. D.3. These fluctuations can lead to 0D carrier confinement (interfacial quantum dots) [96G1, 00L1].

A top-down approach to create quantum dots is again the micro- or nano-structuring of a quantum well by lithography and etching. As for the quantum wires, the lateral dimensions of the dots are usually considerably larger than the thickness of the well, leading to pancake-shaped dots, with the consequence that the quantization energy from lateral confinement is usually small compared to the one from the parent quantum well itself [13L1].

References

[59W1] W.A. Weyl, *Colored Glasses* (Dawson of Pall Mall, London, 1959)
[64W1] R.S. Wagner, W.C. Ellis, Appl. Phys. Lett. **4**, 89 (1964)
[80L1] J.J. LePore, J. Appl. Phys. **51**, 6441 (1980)
[82L1] O. Madelung, U. Rössler (ed.), *Landolt–Börnstein*. New Series, Group III, vol. 17a to i, 22a and b, 41A to D (Springer, Berlin, 1982–2001)
[83Z1] R. Zallen, *Physics of Amorphous Solids* (Wiley, New York, 1983)
[84B1] L.E. Brus, J. Chem. Phys. **80**, 4403 (1984)
[85C1] Y.-C. Chang, L.L. Chang, L. Esaki, Appl. Phys. Lett. **47**, 1324 (1985)
[85E1] A.I. Ekimov, Al.L. Efros, A.A. Onushchenko, Solid State Commun. **56**, 921 (1985)
[90S1] R.A. Stradling, P.C. Klipstein (eds.), *Growth and Characterisation of Semiconductors* (Adam Hilger, Bristol, 1990)
[92B1] U. Becker et al., J. Crystal Growth **125**, 384 (1992)
[93C1] R. Cingolani, R. Rinaldi, Rivista del Nuovo Cimento **16**, 1 (1993)
[93G1] M. Grün et al., Opt. Mater. **2**, 163 (1993); J. Crystal Growth **141**, 68 (1994)
[93M1] C.B. Murray, D.J. Norris, M.G. Bawendi, J. Am. Chem. Soc. **115**, 8706 (1993)
[94G1] M. Grün et al., J. Crystal Growth **138**, 150 (1994)
[95W1] U. Woggon, S.V. Gaponenko, Phys. Status Solidi (b) **189**, 285 (1995)
[96A1] A.P. Alivisatos, Science **271**, 933 (1996)
[96G1] D. Gammon, E.S. Snow, B.V. Shanabrook, D.S. Katzer, D. Park, Phys. Rev. Lett. **76**, 3005 (1996)
[96I1] T. Inoshita, H. Sakaki, J. Appl. Phys. **79**, 269 (1996)

[96K1] H. Kalt, *Optical Properties of III–V Semiconductors: The Influence of Multi-Valley Band Structures*, Springer Series in Solid-State Sciences, vol. 120 (Springer, Berlin, 1996)
[96S1] T. Someya, H. Akiyama, H. Sakaki, J. Appl. Phys. **79**, 2522 (1996)
[97H1] M. Hetterich et al., Phys. Rev. B **56**, 12369 (1997)
[97H2] A. Hartmann et al., Appl. Phys. Lett. **71**, 1314 (1997)
[97R1] M. Rabe et al., Phys. Status Solidi (b) **202**, 817 (1997)
[97W1] U. Woggon et al., Appl. Phys. Lett. **71**, 377 (1997)
[97W2] U. Woggon, *Optical Properties of Semiconductor Quantum Dots*. Springer Tracts in Modern Physics, vol. 136 (Springer, Berlin, 1997)
[98G1] M. Grün et al., Appl. Phys. Lett. **73**, 1343 (1998)
[98L1] D. Lüerßen et al., Phys. Rev. B **57**, 1631 (1998)
[99B1] D. Bimberg, N. Grundmann, N.N. Ledentsov, *Quantum Dot Heterostructures* (Wiley, Chicester, 1999)
[99B2] G. Biasiol, E. Kapon, J. Crystal Growth **201/202**, 62 (1999)
[00B1] G. Biasiol, K. Leifer, E. Kapon, Phys. Rev. B **61**, 7223 (2000)
[99G1] S.V. Gaponenko et al., IEEE J. Lightwave Techn. **17**, 2128 (1999)
[01H1] M.H. Huang et al., Adv. Mater. **13**, 113 (2001)
[00K1] E. Kurtz et al., Thin Solid Films **367**, 68 (2000); ibid. **412**, 89 (2002); J. Crystal Growth **214/215**, 712 (2000); Phys. Status Solidi (b) **229**, 519 (2001)
[00K2] H. Kissel et al., Phys. Rev. B **62**, 7213 (2000)
[00L1] D. Lüerßen, R. Bleher, H. Kalt, Phys. Rev. B **61**, 15812 (2000)
[00N1] S. Nakamura, S. Pearton, G. Fasol, *The Blue Laser Diode: The Complete Story* (Springer, Heidelberg, 2000)
[01K1] C. Klingshirn in Landolt-Börnstein III/34: *Semiconductor Quantum Structures*, Subvol. C *Optical Properties*, Part 1, (Springer, Berlin, 2001)
[01L1] C. Klingshirn (ed.), *Landolt–Börnstein*. New Series, Group III, vol. 34C, Parts 1 and 2 (Springer, Berlin, 2001); (2004); Part 3, ed. by E. Kasper, C. Klingshirn (2007)
[01L2] F. Leiter et al., Physica B **308–310**, 908 (2001)
[01S1] R.L. Sellin et al., Appl. Phys. Lett. **78**, 1207 (2001)
[01V1] I. Vurgaftman, J.R. Meyer, L.R. Ram-Mohan, J. Appl. Phys. **89**, 5815 (2001)
[02K1] E. Kapon, in *Spectroscopy of Systems with Spatially Confined Structures* (2001). NATO Science Series II, vol. 90 (Kluwer, Dordrecht, 2002), p. 243
[02K2] D. Katz et al., Phys. Rev. Lett. **89**, 086801 (2002)
[03M1] T. Mokari, U. Banin, Chem. Mater. **15**, 3955 (2003)
[03R1] A. Rosenauer, *Transmission Electron Microscopy of Semiconductor Nanostructures: An Analysis of Composition and Strain State*. Springer Tracts in Modern Physics, vol. 182 (Springer, Berlin, 2003)
[03T1] A. Tsukazaki, Appl. Phys. Lett. **83**, 278 (2003)
[04C1] N.I. Cade et al., Phys. Rev. B **70**, 195308 (2004)

[04K1] H. Kalt in Landolt-Börnstein III/34: *Semiconductor Quantum Structures*, Subvol. C *Optical Properties*, Part 2, (Springer, Berlin, 2004)
[04L1] J. Lao et al., Appl. Phys. A **78**, 539 (2004)
[04S1] V.A. Scubin, N.N. Ledentsov, D. Bimberg, *Epitaxy of Nanostructures* (Springer, Heidelberg, 2004)
[05N1] K.S. Novoselo et al., PNAS **102**, 10451 (2005)
[06F1] H.J. Fan et al., Small **2**, 561 (2006)
[07G1] A.K. Geim, K.S. Novoselov, Nat. Mater. **6**, 183 (2007)
[07K1] C. Klingshirn, Phys. Status Solidi B **244**, 3027 (2007)
[07L1] E. Kasper, C. Klingshirn (eds.), *Landolt-Börnstein, New Series, Group III*, vol. 34C3 (Springer, Heidelberg, 2007)
[07N1] C. Neumann et al., Phys. Status Solidi B **244**, 1451 (2007)
[07R1] A. Reiser et al., J. Appl Phys. **101**, 054319 (2007)
[07Z1] H.J. Zhou et al., Appl. Phys. Lett. **91**, 181112 (2007)
[08K1] M. Kuno, Phys. Chem. Chem. Phys. **10**, 620 (2008)
[08L1] M. Lorenz et al., Phys. Status Solidi C **5**, 3280 (2008)
[08O1] G.A. Ozin, A. Arsenault, L. Cademartiri, *Nanochemistry: A Chemical Approach to Nanomaterials: Edition 2*, (Royal Society of Chemistry, 2008)
[08P1] J. Planelles-Aragó et al., J. Mater. Chem. **18**, 5193 (2008)
[08S1] Th. Seyller et al., Phys. Status Solidi B **245**, 1436 (2008)
[08Z1] H.J. Zhou et al., Appl. Phys. Lett. **92**, 132112 (2008)
[09N1] A.H.C. Neto, Physics **2**, 30 (2009)
[09S1] V. Schmidt, J.V. Wittemann, S. Senz, U. Gösele, Adv. Mater. **21**, 2681 (2009)
[09W1] M. Willander et al., Nanotechnology **20**, 332001 (2009)
[09Y1] R. Yan, D. Gargas, P. Yang, Nat. Photon. **3**, 569 (2009)
[10A1] M. Grundmann (ed.), Architecture of nano- and microdimensional building blocks. Phys. Status Solidi B **247**, 1275–1392 (2010)
[10B1] F. Bonaccorso, Z. Sun, T. Hasan, A.C. Ferrari, Nat. Photon. **4**, 611 (2010)
[10G1] M. Grundmann (ed.), Architecture of nano- and micro-dimensional building blocks. Phys. Status Solidi B **247**, 1275–1392 (2010)
[10H1] J.A. Hollingsworth, V.I. Klimov, in *Nanocrystal Quantum Dots 2nd edn*, Chap. 1 (CRC, Boca Raton, 2010)
[10K1] C. Klingshirn et al., Phys. Status Solidi B **247**, 1424 (2010)
[10K2] C. Klingshirn, B.K. Meyer, A. Waag, A. Hoffmann, J. Geurts, *Zinc Oxide: From Fundamental Properties towards Novel Applications*. Springer Series in Materials Science, vol. 120 (Springer, Berlin, 2010)
[10L1] appG:10L1 M. Lorenz et al., Phys. Status Solidi B **247**, 1265 (2010)
[10S1] G. Schmidt (ed.), *Nanoparticles: From Theory to Applications*, 2nd edn. (Wiley-VCH, Weinheim, 2010)
[11H1] M. Helfrich et al., J. Cryst. Growth **323**, 187 (2011)
[12N1] K.S. Novoselov et al., Nature **490**, 192 (2012)
[12W1] Q.H. Wang et al., Nat. Nanotechnol. **7**, 699 (2012)
[13C1] O. Chen et al., Nat. Mater. **12**, 445 (2013)

[13K1] A.B. Kairdolf et al., Annu. Rev. Anal. Chem. **6**, 143 (2013)
[13L1] C. Klingshirn (ed.), *Landolt–Börnstein, New Series, Group III*, vol. 34A (Springer, Heidelberg, 2013)
[13T1] P. Tonndorf et al., Opt. Express **21**, 4908 (2013)
[14D1] X. Duan et al., Nat. Nanotechnol. **9**, 1024 (2014)
[16N1] K.S. Novoselov et al., Science **353**, aac9439 (2016)
[18C1] modified from https://www.cfn.kit.edu/img/research_a_nano_photonics/A2_5_Fig2.jpg accessed 10.08.2018
[19D1] From http://www.fotonik.dtu.dk/english/research/nanophotonics/nanodev/research/quantum-dots accessed 12.03.2019
[19R1] From http://en.rusnano.com/press-centre/news/88604 accessed 12.03.2019

Index

A
Abc, 123
  Pekar–Hopfield, 124
Absorber material, 232
Absorption, 21, 53, 54, 69
  coefficient, 22, 238
  constant, 22
  induced, 378
  of quantum dots, 438
  of quantum wells, 409
  phonon-assisted, 374
  spectra, 339
  strong, 93
Acceptor, 274
Acoustic wing, 391
Active
  IR, 168
  optically, 154, 167
  Raman, 154, 167, 168
Alloy
  quaterny, 179
  semiconductor, 179, 284, 285
  ternary, 179
Amalgamation type, 180
Amorphous semiconductor, 284
Anderson model, 281
Angular momentum, 519
Anisotropy
  optical, 58
Approximation
  adiabatic, 196
  Born–Oppenheimer, 196
  Debye's, 159
  Einstein's, 159
  one-electron, 197
  tight binding, 219
Artificial atom, 268
Atomic layer epitaxy (ALE), 535
Atomic orbitals, 518
Attenuated total reflection (ATR), 174, 361
Augmented plane wave model (APW), 198

B
Back-scattering
  enhanced, 291
Band
  alignment, 539
  allowed, 201
Band alignment
  type I, 259
  type II, 421
  type II misaligned, 261
  type II staggered, 259
  type III, 261
Band-edge renormalization, 309
Bandgap, 1, 216
Band offset, 258
Bandstructure
  calculations, 198, 215
  electronic, 200
Band-to-band transition, 235
  direct, 238
  indirect, 238
Band warping, 227, 520
Basis, 136
Basis functions, 510
Beam
  extraordinary, 60
  ordinary, 60
Beer's law, 54
Biexciton, 325
  binding energy, 325
Binding energy, 275

© Springer Nature Switzerland AG 2019

H. Kalt and C. F. Klingshirn, *Semiconductor Optics 1*, Graduate Texts in Physics, https://doi.org/10.1007/978-3-030-24152-0

Birefringence, 58, 59, 123, 358, 469
induced, 65
Black-body radiation, 39, 160
Blinking, 440
Bloch function, 518
Bloch oscillation, 206
Bloch theorem, 200, 216
Bloch waves, 199
Bogoliubov transformation, 103
Boltzmann constant, 38
Boltzmann statistics, 321
Bose–Einstein condensate, 321
Bose–Einstein statistics, 30
Boson, 29, 321
Bottle neck, 332, 344
Boundary conditions, 35, 42, 43, 123
additional (abc), 123, 332
Bound-exciton complexes (BEC), 387, 412, 460
multiple, 394
poly-centric, 277
Bowing parameter, 286
Branch
acoustic, 149, 151
optical, 149, 151
Bravais lattices, 137
Brewster's angle, 50
Brillouin scattering, 72, 171
resonant, 355
Brillouin zone, 139, 147, 151, 200
mini, 185, 259
Broadening
homogeneous, 93, 348, 413
inhomogeneous, 93
Burstein–Moss shift, 309

C

Calorimetric absorption spectroscopy (CAS), 353
Cauchy's theorem, 493
Causality, 492
Cavity-QED, 361
Central-cell corrections, 276, 388
Cgs-system, 13
Chain
monatomic, 147
Character table, 508
Characters, 497, 508
Charge density, 12
Chemical Evaluation by Lattice Fringe Analysis (CELFA), 544
Chemical potential, 38, 246
Chemical shift, 388
Chemical vapor deposition, 540
Classification, 499
Clausius–Mosotti equation, 86
Cleaved-edge overgrowth, 267, 541
Closure, 498
Cluster
percolating, 291
Coefficent
reflection, 42
transmission, 42
Coherence, 72
Coherence volume, 116
Collision-broadening, 469
Commutation relation, 29
Commutator, 29
Compatibility relation, 510
Compensation
self-, 275
Conduction band, 199, 201
Conductivity, 1, 18, 209
frequency-dependent, 300
minimum, 285
Confinement
intermediate, 436
lateral, 432
strong, 437, 445
weak, 436
Conjugacy class, 499
Conservation laws, 47, 113
Continuity equation, 13
Continuous group, 518
Coulomb gauge, 28, 73
Coupled oscillators, 105
Coupled quantum wells, 258
Coupling
strong, 73, 101, 329
weak, 73, 101, 330
Critical point, 259, 366
Crystal
biaxial, 58
lattice, 135
structure, 136
uniaxial, 59
Crystal field, 230, 519
Current, 18
Cyclotron energy, 460

D

Damping, 88
Markov, 84
strong, 121

weak, 84
Dangling bonds, 290
Dead layer, 332
De Broglie wavelength, 251
Deep center, 279
Defect electron, 203
Defects, 181, 273
interstitial, 274
point, 181
substitutional, 274
Deformation potential, 211, 474
Degeneracy
accidental, 511
Degenerate occupation, 247
Dense medium, 116
Density of states, 33, 252
effective, 246
joint, 252, 366
of phonons, 157
Diamagnetic shift, 454, 456, 464
Diamond structure, 505
Diatomic chain, 149
Dichroism, 59, 358
magnetic, 66
Dielectric
confinement, 429
displacement, 85
tensor, 58
Dielectric constant, 19, 321
background, 88, 90
static, 90
Dielectric function, 19, 85, 87, 88, 91, 165, 241, 319, 485
k-dependent, 119
of plasmons, 300
Diffusion
spectral, 440
Diluted magnetic semiconductors (DMS), 461
Dipole
approximation, 76, 77
operator, 77, 86
transition
probability, 77
Direct product, 509
Dislocation, 538
Disorder
compositional, 286
diagonal, 281
off-diagonal, 281
Dispersion, 118, 148
anomalous, 119
folding back of, 230
normal, 119
of phonon polaritons, 168
of polaritons, 416
relation, 30, 31, 118, 147, 354
relation horizontal, 140
spatial, 122, 123, 332, 335, 354, 358
Dispersion relation, 103
of biexcitons, 325
of excitons, 317
of free electrons, 198
of phonons, 155
of photons, 31
of plasmons, 299
Distribution function, 33
Donor, 274
acceptor pairs, 277
Doping
δ-, 280
n-type, 275
p-type, 275
high, 302
modulation, 280
Double group, 512
Drude–Lorentz model, 81, 300
Dulong and Petit, law of, 159

E

Effective mass, 206, 363
approximation, 207, 276, 316
concept, 205
density of states -, 230
force -, 208
longitudinal, 229
momentum -, 208
transverse, 229
Effective medium, 85, 116
Efficiency
differential, 353
Eigenfrequency
longitudinal, 22, 90
transverse, 90
Eigenfunction, 510
Eigenstates, 510
Electrical current density, 12
Electric displacement, 12
Electric field strength, 11
Electromagnetic wave
transverse, 101
Electron gas
quasi-two dimensional, 254
Electron–hole pair, 1
Electron–hole plasma (EHP), 351

Electron-phonon interaction, 209
Electrons, 204
  nearly free, 217
Electron system
  low-dimensional, 251
Electro-reflectance (ER), 244
Element, 498
  neutral, 498
Elliott formula, 341
Ellipsometry, 242
Emission
  spontaneous, 70, 78
  stimulated, 70, 78
Empty lattice, 219
Energy band, 199
Energy flux density, 17, 52, 62
Escape depth, 344
Etalon, 56
Ewald–Bloch theorem, 200
EXAFS, 367
Exchange interaction, 439
  analytic part of, 320
  short range, 320
Exchange splitting, 324, 439
Excitation
  virtual, 72
Excitation spectrum, 411
Exciton, 102, 315
  Bohr radius, 317
  bright, 320
  charge transfer (CT), 364
  core, 324, 367
  creation operator, 321
  dark, 320
  formation of, 351
  Frenkel, 316, 319
  heavy hole, 409
  in 2D, 405
  indirect, 418
  light hole, 409
  localization, 397
  ortho, 320, 370
  para, 320, 370
  reduced mass, 317
  Rydberg energy, 317
  saddle-point, 367
  singlet, 320
  spatially indirect, 418
  translational mass, 317
  triplet, 320
  Wannier, 316, 318
  wave vector, 317
  yellow series, 370
Exciton binding energy, 319, 323
  for effective dimensionality, 407
  in 2D, 406
  in quantum dots, 434
  in quantum wires, 434
Excitonic enhancement, 341
Excitonic molecule, 325
Exciton–LO phonon coupling, 345
Exciton polariton, 329
  in quantum wells, 416
Experimental technique
  μ-Photoluminescence (μ-PL), 443
  attenuated total reflectance (ATR), 361
  ellipsometry, 242
  Fourier transform infrared spectroscopy (FTIR), 165
  hyper Raman scattering (HRS), 359
  integrating sphere, 351
  luminescence yield, 351
  modulation spectroscopy, 243
  photoluminescence excitation spectroscopy (PLE), 425
  Raman and Brillouin scattering, 168
  scanning nearfield optical microscopy (SNOM), 443
  two-photon absorption (TPA), 359
Extended Hückel method, 199
Extinction, 21, 53

F

Fabry–Perot mode, 339, 354
Fabry–Perot resonator, 56
Fano resonance, 371
Faraday configuration, 466
Faraday effect, 66
Faraday rotation, 463
Fermat's principle, 46
Fermi-edge singularity, 309
Fermi energy, 38, 208, 246
Fermi integral, 246
Fermi sea, 299
Fermi's golden rule, 74
Fermi statistics
  correction to, 276
Fermions, 518
Field ionization, 469
Finesse, 57
Finite damping, 106, 120
Four-wave mixing (FWM), 357
Fourier transformation, 126
Fourier transform infrared spectroscopy (FTIR), 165, 166

Fourier transformation, 492
Fröhlich coupling, 210
Fröhlich interaction, 212
Franz–Keldysh effect, 468
Franz–Keldysh oscillations, 470
Free spectral range, 57
Free-to-bound transition, 278
Frenkel excitons, 363
Frequency
  angular, 15
  complex, 126
Fresnel's formulae, 48

G
Gap
  direct-gap, 203
  forbidden, 201
  indirect-gap, 203
  mobility, 287
  optical, 287
Gauss' theorem, 43
g-factor, 455, 462
Giant Rydberg excitons, 372
Glide reflection, 504
Graphene, 262, 264
  luminescence from, 264
  transmittance of, 264
Group, 498
  Abelian, 499
  abstract, 500
  of $\boldsymbol{k}$, 520
  of H, 510
  point, 136
  space, 137, 503
  theory, 497
  translational, 136
Group velocity, 15, 22, 111, 205, 217, 319, 345
  negative, 120
Growth mode
  Frank-Van der Merwe, 538
  Stranski-Krastanov, 269, 538, 544

H
Hamilton operator, 510
Hamiltonian
  of a semiconductor, 195
Hamilton operator, 29
Harmonic oscillator, 29, 153
Hayne's rule, 388
Heavily doped semiconductors, 284
Heavy hole, 227
Helmholtz equation, 14
Helmholtz–Ketteler formula, 84, 87
Hermann–Mauguin symbols, 504
Heteroepitaxy, 537
Highest occupied molecular orbital (HOMO), 363
Hole, 203
  heavy, 222, 257
  light, 222, 257
Homoepitaxy, 537
Homomorphism, 500
Honeycomb lattice, 262
Huang–Rhys factor, 279, 350, 439

I
Identity, 503
Identity Operation or Mapping, 498
Impact ionization, 469
Impurity band, 285
Index of refraction, 485
  complex, 20, 91, 165
Indicatrix, 62, 64
Insulator, 202
  inorganic, 365
Integrating sphere, 352
Intensity, 17
Interaction
  anharmonic, 155
  electromagnetic, 196
Interfaces fluctuations, 539
Intermittency, 440
Intra excitonic transition, 351
Intrinsic carrier density, 247
Invariant expansion, 516
Inversion, 504
Isoelectronic trap, 274, 279
Isometry, 503
Isomorphism, 500

J
Joint density of states, 237

K
Kerr effect, 469
Kerr rotation, 463
k-linear term, 230, 333, 522
Kramers degeneracy, 521
Kramers–Heisenberg dielectric function, 87
Kramers–Heisenberg formula, 84
Kramers–Kronig relation, 113, 165, 241, 469

Kronig–Penney potential, 281

L
Lambert–Beer's law, 53
Landau level, 454, 460
Landau quantization, 464
Landau-level fan, 463
Laplace operator, 12
Lattice
  empty, 218
  reciprocal, 138
Laue condition, 148
Law
  Associative, 498
  Commutative, 499
  of reflection, 45
  of refraction, 45
Level repulsion, 98, 224, 415
Lifshitz model, 282
Light
  cone, 413
  propagation, 91, 102, 115
Light hole, 227
Linear-chain model, 146
Linear combination of atomic orbitals (LCAO), 199, 200, 219
Linear optics, 19
Lineshape
  Lorentzian, 89
  Voigt, 93
Liquid-phase exfoliation, 540
Local field, 86
Localization, 414
  length, 283
  weak, 290
Longitudinal branch, 107
Longitudinal-transverse splitting energy, 90
LO-phonon replica, 345, 347
Lorentz force, 13
Lorentz oscillator, 110, 300
Lorenz–Lorentz equation, 86
Loss function, 165
Lowest unoccupied molecular orbital (LUMO), 363
Luminescence, 344
Luminescence yield, 351, 377
Luttinger parameters, 227
Lyddane–Sachs–Teller relation, 90
Lyman series, 377

M
Magnetic field strength, 12
Magnetic flux density, 12
Magnetic induction, 12
Magnetic moment
  of excitons, 455
Magnetic permeability, 20
Magnetization density, 12
Magneto-polaritons, 458
Mahan exciton, 309
Main section, 60
Mass-action law, 247, 302
Material
  equations, 13
  inverted, 53
  passive, 53
Maxwell's equations, 11
Metal, 202, 208
Metal-organic chemical vapour deposition (MOCVD), 535
Metal-organic vapour phase epitaxy (MOVPE), 535
Metamaterials, 23
Micro-mechanical cleavage, 540
Micro-photoluminescence, 444
Migration enhanced epitaxy (MEE), 535
Miniband, 259, 420
Mixed-mode state, 369
Mobility, 209
Mobility edge, 283, 293, 397
Mode
  evanescent, 122
  normal, 153
  propagating, 122
Modulus
  elasticity, 143
  shear, 144
Molecular beam epitaxy (MBE), 535
Momentum
  conservation, 114, 138
  operator, 76
  quasi, 205
Momentum density, 17, 62
Monolayer Semiconductors, 540
Monolayer semiconductors, 262, 427
Mott transition, 285
Multi-phonon resonance, 165
Multiple quantum well (MQW), 258
Multiplication, 498
  table, 502

N
Nabla-operator, 12
Nano-wires, 541

Nanorods, 267, 432, 541
Na vapor, 110
Nearly free electron model (NFE), 198, 200, 215
Near-zero index materials, 23
Neutron scattering, 155
Nipi, 261, 426
Noether's theorem, 47, 497
Non-crossing rule, 105, 216, 304
Non-degenerate occupation, 246
Non-equilibrium distribution, 248
Non-locality, 491
Normalization in a box, 34
$N \pm 1$ particle problem, 203

O
One-photon transitions, 514
Operator
  annihilation, 29, 154
  creation, 29, 154
  number, 29, 154
Optical activity, 65
Optical anisotropy, 58
Optical energy band, 237
Optical excitations, 81
Optical gain, 78
Optical gap, 343
Optical tunnel effect, 46
Orbital
  atomic, 199
Order, 499
Ordering
  antiferromagnetic, 462
  ferromagnetic, 462
Orthogonalized plane wave model (OPW), 198
Oscillation period, 15
Oscillator
  harmonic, 154
  Lorentz, 81, 84
  model, 81
  Selmaier, 84
Oscillator strength, 86, 87, 319
  giant, 396
Oxidation
  anodic, 447

P
Pair excitation, 299
Parity, 514
Parity forbidden transitions, 370
Particle-wave dualism, 29
Percolation, 291, 399
Periodic table of the elements, 1
Permeability of vacuum, 12
Permittivity of vacuum, 12
Persistent mode, 180
Perturbation theory, 72, 74, 235, 515
  $\boldsymbol{k} \cdot \boldsymbol{p}$, 223
Phase velocity, 15, 22
Phonon, 102, 122, 153, 154, 197, 363
  back-folded acoustic, 185
  back-folded optical, 186
  confined optical, 187
  decay, 155
  dephasing, 174
  dynamics, 173
  interface, 189
  localized, 181
  optical, 154, 163
  replica, 278, 398
  surface acoustic, 127
Phonon polariton, 163
  coherent, 173
Phonon scattering
  inelastic, 212
Phonoriton, 350
Photoelectric effect, 27
Photoluminescence, 53
Photoluminescence excitation spectroscopy (PLE), 425
Photon, 27, 29, 101
  momentum, 30
  polarization, 30
  spin, 30
Photonics, 27
Photovoltage excitation spectroscopy, 470
Piezo coupling, 213
Planck's law, 39, 160
Plane harmonic wave, 15
  in matter, 20
Plane-parallel slab, 54, 95
Plasma, 485
  electron–hole, 321
  oscillations, 298
Plasmon, 102, 297
  intersubband, 307
  intrasubband, 307
  surface, 127
  valence-band, 307
Plasmon-phonon mixed state, 304
Pockels effect, 469
Γ-point, 141, 200
Point defect, 273
Point group, 504, 513

Poisson equation, 28
Polariton, 73, 102, 329, 364, 420
  cavity, 361, 417
  dispersion, 487
  equation, 20, 104, 354
  exciton, 330, 344, 416
  mixed-mode, 107, 337
  operator, 331
  phonon, 163
  propagation, 102
  surface, 23, 127, 129
  surface plasmon, 303
  valley, 430
Polariton branch
  lower, 104, 120
  upper, 104, 120
Polarizability, 85
Polarization, 324
  density, 12, 85, 101
  orientational, 485
  wave, 102
Polaron, 322
  large, 209
  magnetic, 211, 463
  small, 209
Porous silicon, 447
Positron, 204
Potential
  Born–Mayer, 145
  electrostatic, 28
  harmonic, 145
  Lennard–Jones, 145
  Vector, 28
Poynting vector, 17, 61
Precursors, 536
Primitive basis, 503
Principle angle of incidence, 51
Product, 498
  outer or direct, 506
Propagation
  of light, 71, 487
  superluminal, 106
Propagation quantum beats, 111, 360
Pseudo-potential, 198

Q
Quantization energy
  of quantum wires, 266
Quantization of light, 27
Quantization regimes, 436
Quantum
  well, 254
  wire, 266
Quantum confinement, 447
Quantum dot, 190, 268, 413
  colloidal, 268
  electrostatically defined, 269
  fine-structure, 445
  interfacial (IQD), 269
  self-assembled (SAQD), 269, 538, 544
Quantum efficiency, 440
Quantum island, 413
Quantum number, 519
Quantum optics, 27
Quantum well, 517
  coupled, 418
  double, 418
  type II, 418
Quantum wire, 432
  colloidal, 267
  ridge, 267
  superlattice, 541
  T-shaped, 267, 434
  V groove, 541
  V-groove, 267, 434
Quasi-Fermi energy, 248
Quasi-momentum, 141, 154, 205
Quasi-particle, 141, 153, 154, 204

R
Rabi flopping, 78
Rabi splitting, 106
Raman scattering, 72, 154, 168, 170
  two-photon, 358
Rayleigh scattering, 117
Recombination
  non-radiative, 70
  radiative, 70
Recombination center, 274, 279
Reflection, 54, 504
  internal, 344
  total
    attenuated, 46
    frustrated, 46
    internal, 45, 51
  total internal, 94, 344
Reflection coefficient, 48
Reflection spectra, 165, 337
  excitonic, 415
Reflectivity, 43, 49, 93, 485
Refractive index, 21, 71
  complex, 105, 241
Representation, 497, 506
  equivalent, 507

irreducible, 506, 510
reducible, 506
trivial, 507
Resonance
close-lying, 97
isolated, 88
two close-lying, 109
Resonance flourescence, 70, 115
Response function, 491
linear, 19
Reststrahlen band, 95, 104, 165, 332
Retardation, 492
Rotation, 503
Rotoinversions, 504
Rydberg energy, 276

S
Satellites
LO-phonon, 345
Scanning near-field optical microscope (SNOM), 445
Scattering
anti-Stokes, 170
Brillouin, 53
deformation-potential, 212
Mie, 54
Raman, 53
Rayleigh, 53
Stokes, 170
Schrödinger equation, 195
Screw rotations, 503
Second quantization, 28, 102, 321
for phonons, 154
Selection rules, 411, 497, 512
Self-trapping, 363
Semi-classical treatment of radiation, 73
Semiconductor, 202
amorphous, 289
compound, 1
diluted magnetic, 461
direct-gap, 203, 229
element, 1
indirect-gap, 203, 229
multi-valley, 229
narrow-gap, 202, 231
organic, 4
quaternary, 3
semi-magnetic, 461
single-valley, 229
ternary, 3
wide-gap, 202
Semimagnetic semiconductors (SMSC), 461
Semimetal, 202
Shake-up processes, 310
Shift
chemical, 276
Single quantum well (SQW), 258
Singlet, 365
Singlet-triplet splittings, 363
SI system, 13
Slow light, 110, 113
Small damping, 91
Snell's law, 45
Sommerfeld enhancement, 321, 366
Sommerfeld factor, 407
Sp^3 binding, 519
Space
real, 138
reciprocal, 138
Space group, 512
non-symmorphic, 513
symmorphic, 512
Spatial dispersion, 129, 492
Specific heat, 159
Spectroscopy
$\boldsymbol{k}$-space, 354
modulation, 243
momentum-space, 354
single dot, 440
spatially resolved, 413
Spectrum
absorption, 487
luminescence, 487
reflection, 487
Speed of light, 15
Spherical rotation group, 518
Spin, 30, 512, 516
Spin glass, 462
Spin-orbit coupling, 230, 262
Spin-orbit interaction, 427
Spin-orbit splitting, 516, 519
Spin-valley locking, 263
Splitting of states, 512
Square-well potential, 255
S-shaped temperature dependence, 399, 414
Star of $\boldsymbol{k}$, 519
Stark effect, 468
excitonic, 456
quantum-confined, 441, 472
State
continuum, 256
localized, 540
quantized, 256
Statistics
Boltzmann, 38

Bose–Einstein, 38
Fermi–Dirac, 38, 245
of phonons, 157
Stokes shift, 411
Stokes' theorem, 43
Stop band, 94, 104, 122, 163
Strain, 339
biaxial, 475, 477
uniaxial, 474
String
one-dimensional, 143
Strong-field limit, 454
Subband, 252, 255, 256
heavy-hole, 256
light-hole, 256
quasi-one-dimensional, 266
Subband transitions, 409
Subgroup, 499
Superlattice
corrugated, 539
free-standing, 538
quantum wire, 541
strained-layer, 538
Superlattice (SL), 184, 259, 420
doping, 261, 426
free-standing, 478
short-period, 422, 424
Stark-, 422
strain-symmetrized, 478
strained, 478
strained-layer, 184, 422
type-II, 421
Surface charge density, 297
Surface layer
exciton-free, 332
Surface phonon polariton, 174
Surface plasmon, 302
Susceptibility, 19, 85
Symmetry, 58
group, 501
operations, 501
translational, 114
System
four level, 78
three level, 78

T

Tailing parameter, 286, 288
Tauc regime, 287, 343
Thickness
critical, 538
Thin-prism method, 357
Three-photon absorption, 360
Tight-binding approximation, 199
Time of flight, 358, 417
Time reversal, 521
Total internal reflection
attenuated, 361
Total reflection
attenuated, 46
Transition
free-to-bound, 274
internal, 280
intra-excitonic, 378
matrix element, 514
probability, 75, 347
rate, 235
Transition-metal dichalcogenide (TMD), 262, 427
Translation, 503
group, 512
invariance, 138
vector, 135
Transmission, 54, 95
Transmission coefficient, 48
Transmittivity, 43, 49
Trion, 325
Triplet, 365
Triplet fluorescence, 365
Two-electron transitions, 392
Two-hole transition, 393
Two-level system, 69
Two-particle excitation, 299
Two-photon absorption (TPA), 358, 369
Two-photon spectroscopy, 358
Two-photon transitions, 515
Two-polariton state, 359
Two-photon Raman scattering (TPRS), 358

U

Ulbricht-sphere, 352
Umklapp process, 157
Uncertainty principle, 70
Uncertainty relation, 126
Unit cell, 135
Urbach–Martienssen rule, 342
Urbach tail, 287, 348

V

Vacancies, 275
Vacuum state, 315, 378
Valence band, 199, 201
heavy-hole, 521
light-hole, 521

Valley degree of freedom, 263, 427
Valleytronics, 263
Van Hove singularity, 240
Vector potential, 73, 75
Velocity
  front, 107
  group, 106, 107, 144, 147, 358
  negative group, 107
  phase, 71, 107, 144, 147
Vibrons, 363
Virtual crystal approach, 179, 285
Virtual excitation, 70
Voigt configuration, 466
Vollmer-Weber growth mode, 538

W
Wannier function, 205
Wannier–Stark ladder, 473
Warping, 230
Wave
  back-scattered, 217
  evanescent, 45, 361
  longitudinal, 22, 144
  number, 31
  packet, 16
  polarization, 23
  transverse, 16, 144
  vector, 15, 32
Wave equation
  harmonic, 144
  in a dielectric medium, 20
  in vacuum, 14
Wave functions
  product, 516
Wavelength, 15, 21
Weak field limit, 454
Well-width fluctuation, 284, 288
Wigner–Seitz approach, 221
Wigner–Seitz cell, 140
Wurtzite structure, 505

X
XANES, 367

Z
Zeeman splitting, 454, 456, 464
  giant, 462
Zero-phonon line, 395
Zero-point energy, 29
Zincblende structure, 505